Incomplete concerted evolution in the non-hybrid diploid *Clematis fremontii* S. Watson (Ranunculaceae)

A Thesis

Presented for the

Master of Science Degree

The University of Tennessee at Chattanooga

Meredith C. Montgomery

December 2009

Acknowledgements

Dr. Joey Shaw served as my committee chair, and Dr. Margaret Kovach and Dr. Stylianos Chatzimanolis served on my committee. Dr. Hill Craddock managed the *Clematis fremontii* plants in the UTC greenhouse. Stacy Huskins assisted significantly in the collection of *C. fremontii* from Missouri and Kansas. Clea Klagstad assisted in lab work associated with cloning. Several individuals assisted in the collection of plant material, including J. Shaw, R. Ware, M. Adams, D. Estes, J. Evans and G. Woodworth. The Department of Biological and Environmental Sciences at the University of Tennessee at Chattanooga supported me as a graduate student. The Marie Mellinger Research Grant provided by the Georgia Botanical Society, UTC Faculty Start-up Funds for J. Shaw, the UTC Graduate School and Provost award, and awards from the Geospatial Information Technology Association funded my research. I thank J. Shaw for serving as a guide through my long and winding road of research and for pushing me to reach my potential. I thank fellow and former graduate students S. Huskins, M. Binkley, and I. Cohen for their invaluable friendship. Finally, I thank my family for their love and support.

Abstract

Clematis fremontii S. Wats. (Ranunculaceae) is a diploid perennial forb with strict habitat requirements that occurs in isolated populations in the cedar glades of Missouri and the mixed-grass prairies of Kansas and Nebraska. Based on geographical isolation and morphological differences, this species was formerly subdivided into two varieties, the autonymic prairie variety and the glade variety, *C. fremontii* var. *riehlii* Erickson. Interestingly, two disjunct populations of *C. fremontii* inhabit cedar glades have been recently located in Rome, GA and Chattanooga, TN. Initially, the focus of my research was to gain insight into whether the two eastern populations are recent introductions or disjunct relict populations.

The discovery of these populations and North American *Clematis* taxonomy is the subject of Part 1 of this thesis. The second part of this thesis describes a survey of highly variable chloroplast DNA regions that led to the conclusion that chloroplast DNA does not contain enough variability to distinguish populations of *C. fremontii*. Because previous studies showed that nuclear ribosomal (ITS) sequences were variable at the population level in *C. fremontii*, sequences from this nrDNA gene region were generated directly from PCR products. The third part is the centerpiece of my research and includes a description of the heterozygosity observed among directly sequenced PCR products, the uncovering of pseudogenes, the myriad alleles uncovered through a massive cloning effort, and is a cautionary description about the shortcomings of this gene region, especially for studies below the level of subgenus. Unexpectedly, it appears that a couple of paralogous pseudogenes may allude to the eastern populations being related to a population in Missouri. Since nrDNA ITS sequences proved inadequate for revealing population relatedness among populations of *C. fremontii*, the fourth part is a preliminary investigation of the same populations using single-copy nuclear gene glyceraldehyde 3-phosphate dehydrogenase (*G3pdh*) sequence data. While the data set for this gene region is incomplete, it appears that *G3pdh* is single copy in *Clematis fremontii* and variable enough to get at the original goals of this thesis project. The last part of this thesis is an attempt to focus a conservation effort on the Tennessee plants, located entirely on private and commercially zoned property, so that the majority of the genetic diversity of this population can be transferred to a protected site on state lands.

Table of Contents

List of Tables

List of Figures

Part 1

Part 3

Part 4

Part 1: **Overview of *Clematis fremontii* S. Watson (Ranunculaceae)**

INTRODUCTION

Comprised of about 280 species, *Clematis* L. has a worldwide distribution and is one of the largest genera in Ranunculaceae (Yang and Moore 1999). Ranunculaceae is considered by most botanists to be among the most basal groups of eudicots (APG II, 2003) and is primarily characterized by herbaceous plants with a multi-carpelled gynoecium. Within Ranunculaceae, *Clematis* is characterized by single-seeded achenes, numerous stamens, and an apetaloid corolla. *Clematis* is used widely in horticulture and is among the most popular ornamental plants. Species of this genus used most frequently in gardens are those with a viny growth habit (Grey-Wilson 2000). Several cultivars exist; however numerous wild forms such as *C. alpina* (L.) Mill., *C. macropetala* Ledeb., *C. montana* Buch.-Ham. ex DC., and *C. tangutica* (Maxim.) Korsh. are also grown for horticultural use (Grey-Wilson 2000). Other documented anthropogenic uses of *Clematis* include European use of crushed leaves as a vesicant, use of *C. hirsutissima* Pursh as a horse stimulant and as snuff by Northwestern Native Americans, and medicinal use of *C. baldwinii* Torr. & A. Gray by the Seminole peoples of Florida (Erickson 1943). Chromosomal counts of 42 different *Clematis* species revealed that *Clematis* is a diploid genus with eight pairs of chromosomes (2n=16) (Gregory 1941). Polyploidy is rare in *Clematis* but has been documented in three hybrid cultivars (Erickson 1943). To date, polyploidy has not been observed in wild uncultivated *Clematis*.

Tremendous morphological diversity exists within *Clematis*, which is reflected in the varied infrageneric classification systems derived from conventional morphological treatments (Erickson 1943; Keener & Dennis 1982; Tamura 1987; Johnson 1997; Yang and Moore 1999; Grey-Wilson 2000). The most recent classification, divides *Clematis* into nine subgenera: *Flammula* DC.; *Clematis; Cheiropsis* (DC.) Peterm.; *Pseudoanemone* (Prantl) Grey-Wilson; *Tubulosae* (Decne) Grey-Wilson; *Atragene* (L.) Torrey & Gray; *Archiclematis* (Tamura) Grey-Wilson; *Campanella* Tamura; and *Viorna* Tamura, non Reichb. (Grey-Wilson 2000). The genus *Clematis* mostly has a Eurasian distribution with a few exceptions. Subgenus *Pseudoanemone* occupy Africa, and subgenus *Viorna* largely occupies North America. Subgenera *Atragene* and *Flammula* both have a holarctic distribution.

Irresolute systematic treatment also plagues the *Viorna* group, which has recently been revised based on morphological, anatomical and palynological data (Yang and Moore 1999). Subgenus *Viorna* is characterized by erect sepals and hirsute stamens and, with the exclusion of

two east Asian species and one Eurasian species, subg. *Viorna* is a North American group (Grey-Wilson 2000). Erickson's (1943) treatment is the most widely accepted classification of North American *Clematis* sect. *Viorna*, and he divided this section into five subsections. Since Erickson's (1943) treatment, sect. *Viorna* has been raised to subgeneric rank (Keener & Dennis 1982, Tamura 1987, Snoeijer 1992, Yang and Moore 1999, and Grey-Wilson 2000). Unfortunately, Erickson's subsectional names lacked Latin diagnoses and therefore have no nomenclatural standing, however several botanists have referred to Erickson's subsectional classification with little refute (Keener & Dennis 1982, Yang & Moore 1999, Grey-Wilson 2000). Erickson (1943) placed subsections *Euviornae* (woody vines), *Viticellae* (*C. crispa* L.), *Baldwinianae* (*C. baldwinii*), *Hirsutissimae* (erect herbs with pinnate leaves), and *Integrifoliae* (erect suffrutescent herbs with simple leaves) within subgenus *Viorna*. Yang and Moore (1999), however, suggested separating subgenus *Viorna* into four subgenera: *Campanella*, *Tubulosa*, *Viorna*, and *Integrifolia*. The classification derived by Yang and Moore (1999) places Erickson's subsections *Baldwinii*, *Hirsutissimae*, and *Integrifoliae* within subg. *Integrifolia* sens. nov. and *Viticellae* and *Euviornae* within subg. *Viorna*. Thus, the revision of Yang and Moore places only those North American *Clematis* species with an erect habit and *C. integrifolia*, a Eurasian disjunct, into subg. *Integrifolia*. However, in their subgeneric classification of *Clematis* in temperate North America north of Mexico, Keener and Dennis (1982) place erect herbs or suffruticose scandent vines in subgenus *Viorna* A. Gray. Clearly, an accepted classification of *Viorna* sensu lato has yet to be established, as was duly noted by Keener and Dennis; "No subgeneric group of *Clematis* has been subjected to as much taxonomic opinion and nomenclatural confusion as subgenus *Viorna* (1982)." Subsequently, the evolutionary relationships within *Clematis* remain obscure, and much more phylogenetic research is necessary to elucidate these relationships and generate a well-resolved cladistic analysis. Phylogeographic research of *C. fremontii* S. Watson will add to the small pool of existing phylogenetic research of *Clematis*.

Erickson included four species within his subsection *Integrifoliae*, *C. ochroleuca* Aiton, *C. albicoma* Wherry, *C. viticaulis* Steele, and *C. fremontii* sensu lato (1943). He suggested the addition of *C. fremontii* S. Watson var. *riehlii* Erickson based on differences in distribution and leaf morphology. However, C. S. Keener (1967) later determined that extensive overlap in the ranges of variant populations rendered recognition of these varieties impossible. Prior to

Erickson, both *C. albicoma* and *C. fremontii* had been included within *C. ochroleuca* (Gray 1892, Small 1933, James 1883). However, geographic edaphic and morphological differences clearly distinguish the three species. Keener (1967, 1975) raised *C. albicoma* var. *coactilis* to *C. coactilis* (Fernald) Keener, with the suggestion that *C. coactilis* could be a result of hybridization between *C. albicoma* and *C. ochroleuca*.

Members of Erickson's (1943) subsection *Integrifoliae* are limited to narrow edaphic and geographic ranges (Figure 1). The widely distributed *C. ochroleuca* is restricted to mafic substrates of the Piedmont, ranging from northeast Georgia to Staten Island. *Clematis fremontii* is limited to dolomitic or limestone substrates of calcareous glades and prairies in Missouri and Kansas. *Clematis albicoma*, *C. coactilis*, and *C. viticaulis* are known from the Virginias. *Clematis coactilis* is found in shale barrens with rare occurrences on sandstone, dolomite, or limestone outcrops. *Clematis albicoma* and *C. viticaulis* are both restricted to shale barrens of Upper Devonian Brallier Formation. *Clematis albicoma*, *C. viticaulis*, and *C. coactilis* occur separately in isolated populations with no record of sympatry. In summation of subsection *Integrifoliae*, Erickson (1945) noted the narrow geographic range of all species within the subsection, exclusive of *C. ochroleuca*, and suggested this distribution could be relictual in nature. As such, the recent discovery of two populations of *C. fremontii* in southeast Tennessee and northwest Georgia, well outside the established range of this species, was an intriguing surprise.

Clematis fremontii is an edaphic forb with strict habitat requirements, including minimal shade cover and dry, shallow, calcareous soils (Erickson 1945) (Figure 2). It occurs in isolated populations on limestone and dolomitic bedrock in portions of the mixed-grass prairies of Kansas and Nebraska and in dolomitic glades of the Ozark plateau of Missouri (Learn and Schaal 1986). Isolation of populations is likely absolute with virtually no interbreeding or seed dispersal, as populations are separated by several miles, animal dispersal is unlikely and seeds lack the feathery plumes necessary to travel great distances via wind. *Clematis fremontii* is a long-lived perennial and has been characterized as diploid (2n=16) (Erickson 1943, Gregory 1941). It has variable growth habit and leaf morphology, which can range from lanceolate to broadly ovate. This variation is striking enough that Erickson (1943) subdivided this species into a prairie and a glade variety based on differences in leaf morphology and habitat type. Although

these varieties are no longer recognized, the subdivision is a testament to the extent of morphological variation within the species.

Clematis fremontii is of primary interest because two disjunct populations were recently discovered in Hamilton County, Tennessee and Floyd County, Georgia. If these populations are native, rather than introduced, then they effectively represent critically endangered populations in both Tennessee and Georgia, as these are the only two documented occurrences outside of the established range of *C. fremontii*. The populations have striking parallel histories in that both were formerly misidentified. The Floyd County record, deposited to GA in 1939 by Dr. Harold Jones was misidentified as *C. ochroleuca*, and the Hamilton County, TN record deposited to UCHT by Margaret Smith in the 1920s, was labeled *C. ovata* Pursh, a synonym for *C. ochroleuca*. It appears that both of these lost populations of *C. fremontii* were rediscovered in 2003 in Georgia (Chafin and Ware 2003) and 2006 in Tennessee (Horn and Shaw 2007). Therefore, herbarium records place populations of *C. fremontii* in both eastern locations as far back as 70 years ago.

Both locales are small areas of former human inhabitance. The Floyd County population is now protected on the Berry College campus, and thus faces no immediate threat, aside from the encroachment of dense pines and brush upon the plants (Figure 3). However, the Hamilton County population is currently under substantial threat. The cedar glade is located on property that is surrounded by development and zoned for industrial use (Figure 4). Several woody invasive plant species are growing in the glade, which could lead to severe habitat degradation including the loss of direct sunlight, a key glade characteristic. The domestic nature of these two populations, in conjunction with the aesthetic appeal of *C. fremontii*, suggest possible anthropogenic introduction of the species to the Southeast. If the populations were introduced, the extant herbarium records indicate the introductions must have occurred prior to 1939 (GA) and the1920s (TN). Since *Clematis* is commonly used in landscaping and *C. fremontii* has a showy inflorescence, it is possible that the southeastern populations represent recent introductions. However, in a document from the same era as the herbarium records, Erickson (1943) noted that aside from *C. crispa* L. and *C. texensis* Buckley, species of the *Viorna* section of *Clematis* are rarely planted for ornamental purposes (1943). A more recent text written for gardeners and horticulturalists echoes Erickson's sentiments regarding the use of subsect. *Integrifoliae* species in cultivation (Grey-Wilson 2000). In addition, the habitat of *C. fremontii* is

restricted to rocky barrens, rendering the species relatively ill fit for horticultural use (Erickson 1943). The latter information does not disprove the presumption that these populations are introduced, but it does allow the alternative hypothesis that they are naturally occurring relicts. Bolstering the aforementioned hypothesis is the occurrence of several other glade endemics on the Hixson, TN property. To date thirteen glade associates have been inventoried on the property (Table 1). Most notable are *Silphium terebinthinaceum* Jacq., *Manfreda virginica* (L.) Salisb. ex Rose, and *Scutellaria parvula* Michx., which are glade associates found in several of the Midwestern glades that contain *C. fremontii*.

In his treatment of subgenus *Viorna* Erickson (1943) noted, "The species of the subsection *Integrifoliae* are, with the exception of *C. ochroleuca*, very limited in their geographical range, suggesting strongly that they represent relict forms." Exploring the validity of Erickson's statement is now possible through the application of molecular tools and techniques. Surprisingly, despite the size of this genus and the discord regarding its classification system, few molecular studies have examined *Clematis.*

The paucity of molecular research regarding *Clematis* extends through the Ranunculaceae. Molecular research within Ranunculaceae has been dominated by Hoot, Johansson, Jansen, and Schuettpelz. An early chloroplast DNA restriction site mapping study examined cpDNA sequence divergence and phylogenetic relationships among five species of Ranunculaceae, *Caltha palustris* L., *Ranunculus bulbosus* L., *R. fascicularis* Muhl. ex Bigelow, *R. recurvatus* Poir., and *Trollius ledebourii* Rchb. (Johansson and Jansen 1991). Chloroplast DNA sequence divergence was significantly higher in the large and small single copy regions than in the inverted repeat, with total divergence ranging from 0.2% to 9.6%. Johansson and Jansen concluded that restriction site mapping was applicable to intergeneric comparisons within *Ranunculaceae* and interspecific analyses within *Ranunculus*. Hoot and Palmer (1994) used cpDNA cleavage sites for 10 restriction enzymes to map the chloroplast genome of eight genera of *Ranunculaceae*, including two species of *Clematis*. Several inversions were found in the chloroplast genome of *Clematis* species. The plastid *atpB-rbcL* intergenic spacer region and the nuclear ribosomal ITS regions were used to examine the phylogeography of 20 *Anemone* species (Schuettpelz et al. 2002). Both regions were informative at the interspecific level with 62 and 134 parsimoniously informative characters generated by these regions, respectively. The single most parsimonious tree resulting from analysis of the combined regions yielded results almost

identical to those obtained by Hoot and Palmer in 1994 with the use of morphological and plastid restriction site data.

To date only two molecular studies have focused on *Clematis.* Miikeda et al. (2006) conducted a phylogenetic study of the genus. Using only 33 accessions, representing 33 species, they generated two phylogenetic Neighbor Joining trees using ITS1-5.8S-ITS2 sequences of the nuclear ribosomal cistron and combined sequences of five different noncoding cpDNA regions. The ITS phylogeny was as resolved as the five region cpDNA phylogeny, indicating ITS is an informative region for assessing interspecific relationships within *Clematis*. Although the species included in this study represented each subgenus from around the world and six purportedly variable molecular markers were used, few nuclear substitutions were found across taxa. Miikeda et al. (2006) concluded that within *Clematis*, the combination of high morphological diversity and low molecular diversity suggests a recent rapid radiation of the genus.

In the early years of population-level molecular research, Learn and Schaal (1986) used RFLP analysis of the nontranscribed spacer (NTS) of the nuclear DNA ribosomal cistron to assess the genetic diversity of *C. fremontii* in 10x10 meter quatrats of a single population in Victoria Glade, MO. Each quadrat, which consisted of a mean of 12 individuals, displayed 95% of the total genetic diversity for the entire population, indicating future research of *C. fremontii* at the population level should include roughly the same number of individuals per population.

STUDY OBJECTIVES

At the outset of this course of study the main objective was to use molecular tools to assess the origin of the two disjunct southeastern *C. fremontii* populations as well as to investigate the phylogeography of *C. fremontii* in relation to Erickson's (1943) other subsection *Integrifoliae* species within the eastern United States: *C. coactilis, C. albicoma*, *C. viticaulis*, and *C. ochroleuca.* To that end, the chloroplast and nuclear genomes were screened to search for regions displaying intraspecific variation. To verify that the southeastern populations were in fact *Clematis fremontii* and not a separate species, nDNA sequences of these populations were compared to naturally occurring Midwestern populations of *C. fremontii* as well as to other subsection *Integrifoliae* taxa. Although the original intent of this research was to assess the origin of the two disjunct Southeastern *Clematis fremontii* populations, an unusually high

frequency of intraindividual polymorphism within the ITS region prompted an alternate investigation. Subsequent to this discovery, the focus of research became an exploration, through extensive cloning and pseudogene analysis, into the extensive occurrence of intragenomic polymorphism in ITS nrDNA sequences of *C. fremontii* and related taxa within subsect. *Integrifoliae*.

Therefore, Part two of this thesis, entitled "A survey of 5 highly variable chloroplast gene regions reveals low genetic differentiation between *Clematis* subsection *Integrifoliae* species" is aptly named. Part three of this thesis, entitled "Incomplete concerted evolution in the non-hybrid diploid *Clematis fremontii*" was written as a publishable unit. This part describes the ITS gene region in terms of its molecular evolution, importance in systematic biology, population genetics, and phylogeography, and the unusually high frequency of intraindividual polymorphism discovered in *C. fremontii*. The fourth section of this thesis, Part four, entitled "Single Copy Nuclear Gene *G3pdh*: a Promising Population Level Marker for *C. fremontii*" is a preliminary investigation of the same original goals with a low-copy nuclear gene region. In the final part of this thesis includes general conclusions about *C. fremontii* and an outline of future directions.

LIST OF REFERENCES

APG. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. Botanical Journal of the Linnean Society. 141: 339-436.

Chafin and Ware 2003. Georgia Botanical Society Newsletter 77(6).

Erickson, R.O. 1943. Taxonomy of *Clematis* section *Viorna*. Annals of the Missouri Botanical Garden 30(1): 1-62.

Gregory, W.C. 1941. Phylogenetic and Cytological Studies in the Ranunculaceae. Transactions of the American Philosophical Society 31(5): 443-521.

Gray, A. 1892. "Ranunculaceae," pp. 1-57. In: B.L. Robinson (ed.), Synoptical flora of North America, Vol. I, pt. 1. American Book Co., New York.

Grey-Wilson. 2000. Clematis, the genus: a comprehensive guide for gardeners, horticulturists, and botanists.

Hoot, S.B. and J.D. Palmer. 1994. Structural rearrangements, including parallel inversions, within the chloroplast genome of *Anemone* and related genera. Journal of Molecular Evolution. 38(3): 274-281.

Horn, D.D. and J. Shaw. 2007. Noteworthy Collections: Tennessee. Castanea 72(1): 48-49.

James, J.F. 1883. A revision of the genus *Clematis* of the United States; embracing descriptions of all the species, their systematic arrangement, geographical distribution, and synonomy. Journal of Cincinnati Society of Natural History 6: 118-135.

Johansson J. T. and R. K. Jansen. 1991. Chloroplast DNA variation among five species of Ranunculaceae: structure, sequence divergence, and phylogenetic relationships. Pl. Syst. Evol. 178(1-2): 9-25.

Johnson, M. 1997. The genus *Clematis*. Magnus Johnson Plantskola AB.

Keener, C. S. 1967. A biosystematic study of Clematis subsection Integrifoliae (Ranunculaceae). J. Elisha Mitchell Sci. Soc. 83: 1-40.

Keener, C. S. 1975. Studies in the Ranunculaceae of the Southeastern United States. I. *Anemone* L. Castanea 40(1) 36-44.

Keener, C. S. and W. M. Dennis. 1982. The Subgeneric Classification of Clematis (Ranunculaceae) in Temperate North America North of Mexico. Taxon 31(1): 37-44.

Learn, G.H. Jr. and B. A. Schaal. 1986. Population Subdivision for Ribosomal DNA Repeat Variants in Clematis fremontii. Evolution 41(2): 433-438.

Miikeda, O., K. Kita, T. Handa, T. Yukawa. 2006. Phylogenetic relationships of *Clematis* (Ranunculaceae) based on chloroplast and nuclear DNA sequencesBotanical Journal of the Linnean Society 152: 153–168.

Schuettpelz, E., S.B. Hoot, R. Samuel, and F. Ehrendorfer. 2002. Multiple origins of Southern Hemisphere *Anemone* (Ranunculaceae) based on plastid and nuclear sequence data. Plant Systematics and Evolution. 231: 142-151.

Small, J.K. 1933. Manual of the southeastern flora. Publ. by the author, New York.

Snoeijer W. 1992 A suggested classification for the genus *Clematis*. Clematis 1992. 7-20

Tamura, M. 1987. A classification of genus Clematis. Acta Phytotaxonomica et Geobotanica 38: 33–44.

Yang T.Y.A., Moore D.M. 1999. A revision of the Viorna group of species (section Viorna sensu Prantl) in the genus Clematis (Ranunculaceae). Systematics and Geography of Plants 68:281–303.

Table 1. Species list of woody and herbaceous plants located on the Hixson, TN property. Asterisks denote glade associate species.

Woody Species (n=24)	**Herbaceous species** (n=21)
Acer saccharum Marsh.	*Asclepias syriaca* L.
Albizia julibrissin Durazz.	*A. verticillata* L.
Berchemia scandens (Hill) K. Koch *	*Chamaecrista fasciculata* (Michx.) Greene
Bumelia lycioides L.	*Chamaesyce nutans* (Lag.) Small
Carya alba (L.) Nutt.	*Clematis fremontii* S. Watson *
Celtis occidentalis L.	*Croton monanthogynus* Michx. *
Cornus florida L.	*Euphorbia corollata* L. *
Fraxinus sp.	*Eupatorium altissimum* L.
Hedera helix L.	*Fragaria virginiana* Duchesne
Juglans nigra L.	*Gaura filipes* Spach
Juniperus virginiana L. *	*Helianthus hirsutus* Raf.
Ligustrum sinense Lour.	*Hypericum sphaerocarpum* Michx. *
Pinus virginiana Mill.	*Ipomoea pandurata* (L.) G. Mey.
P. taeda L.	*Manfreda virginica* (L.) Salisb. ex Rose *
Prunus serotina Ehrh.	*Physostegia virginiana* (L.) Benth.
Quercus nigra L.	*Potentilla simplex* Michx. *
Quercus phellos L.	*Rudbeckia fulgida* Aiton
Quercus prinus L.	*R. hirta* L. *
Rhamnus caroliniana (Walter) A. Gray	*Ruellia humilis* Nutt. *
Rhus copallinum L.	*Sericocarpus linifolius* (L.) Britton, Sterns & Poggenb.
R. glabra L.	*Scutellaria parvula* Michx. *
Smilax rotundifolia L.	*Silphium terebinthinaceum* Jacq. *
Ulmus alata Michx. *	
Viburnum rhytidophyllum Hemsl.	

Figure 1. Distrubution map of North American taxa included within Erickson's subsection *Integrifoliae*. Members of Erickson's (1943) subsection *Integrifoliae* are limited to narrow edaphic and geographic ranges. *C. ochroleuca* is restricted to mafic substrates of the Piedmont, ranging from northeast Georgia to Staten Island. *Clematis fremontii* is limited to dolomitic or limestone substrates of calcareous glades and prairies in Missouri and Kansas. *Clematis albicoma*, *C. coactilis*, and *C. viticaulis* are known from the Virginias. *Clematis coactilis* is found in shale barrens with rare occurrences on sandstone, dolomite, or limestone outcrops. *Clematis albicoma* and *C. viticaulis* are both restricted to shale barrens of Upper Devonian Brallier Formation.

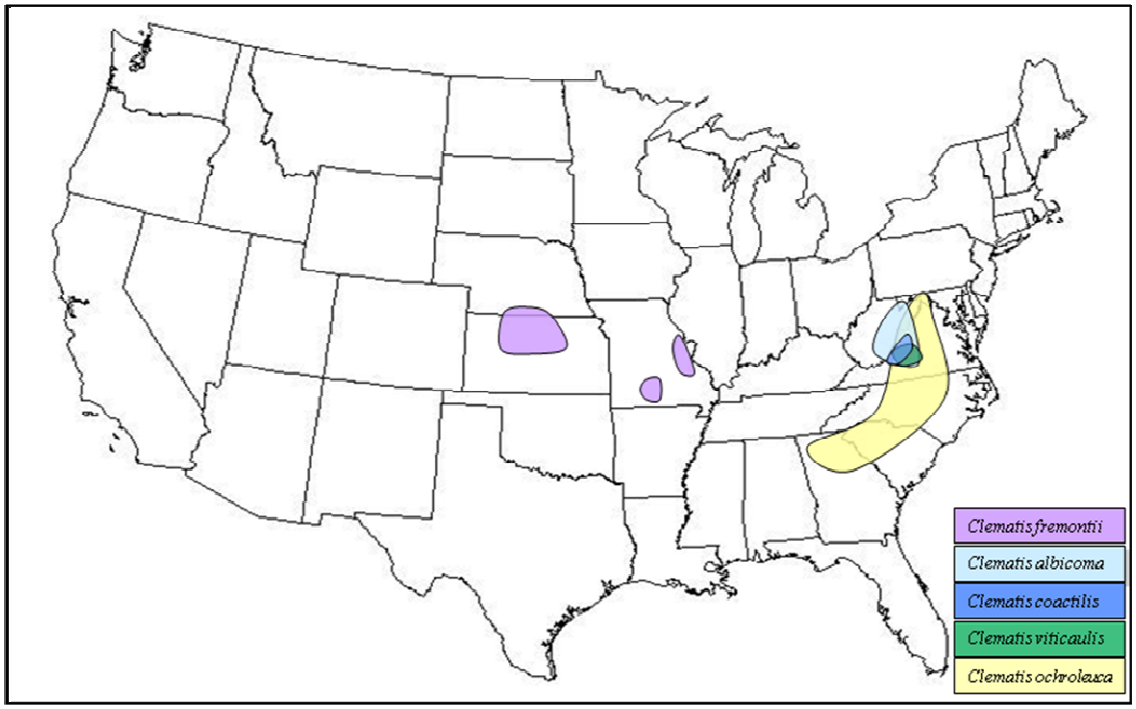


Figure 2. *Clematis fremontii*. *Clematis fremontii* is a long-lived, edaphic, pernnial forb with strict habitat requirements, including minimal shade cover and dry, shallow, calcareous soils (Erickson 1943). It occurs in isolated populations in the mixed-grass prairies of Kansas and Nebraska and in dolomitic glades of the Ozark plateau of Missouri (Learn and Schaal 1986). Two disjunct populations are known from Hamilton County, Tennessee and Floyd County, Georgia. Note the exposed limestone rock in the background.

Figure 3. *Clematis fremontii* population in Rome, Georgia (Floyd County). In 2003, the population was discovered on Berry College campus. Dense pines and brush surround the plants, but the population is monitored and wiring has been constructed around the plants to shield them from wildlife.

Figure 4. Aerial photograph of the 8.75 acre property in Hixson, TN that contains the recently discovered (2006) *C. fremontii* population. Property lines are marked in white. The property is surrounded by developed land and is zoned for industrial use.

Part 2: **Utility of chloroplast DNA in assessing intraspecific variation within *Clematis fremontii***

INTRODUCTION

When using molecular tools to infer phylogenetic or population level relationships among plants, a major limitation is the availability of a molecular marker that will display an adequate level of genetic variability to distinguish study units; i.e., a region of the genome that displays genetic differences unique to each taxon or population.

Although molecular techniques like microsatellites, allozymes, and AFLPs are common tools at low levels of study because they comparatively yield more variation among closely related study units than do sequence data, constructing haplotype networks with these markers is difficult as they cannot be historically ordered (Schaal et al., 2003). A phylogeographic approach to studying the geographical distribution of genealogical lineages (Avise 2000) offers an alternative method of assessing plant populations in that a gene (or several genes) is sequenced for multiple accessions in several populations to reveal their evolutionary relatedness. Through analyzing genealogically ordered sequence data, historical perspective of population-level relationships may be revealed.

Phylogeographic studies of plants largely rely on chloroplast and nuclear DNA sequences to investigate population level relationships. Nuclear DNA (nDNA) typically displays more genetic variability than chloroplast DNA (Small et al. 2004, Sang 2002). However, the cytoplasmically and uniparentally inherited organellar chloroplast genome offers several advantages over biparentally inherited nuclear DNA. Analysis of nDNA is complicated by the possible occurrence of heterozygosity (often requiring cloning to separate alleles), recombination, and paralogous loci, none of which are relevant for cpDNA analysis. Also, because cpDNA is uniparentally inherited and effectively haploid, it offers the advantage of one quarter the amount of relative coalescent time to that of nDNA (Rosenberg and Nordborg 2002). Furthermore, recent studies by Shaw et al. (2005, 2007), which highlighted the most variable or quickly evolving cpDNA regions, have shown cpDNA harbors much more genetic variability than previously thought. However, when choosing which region of the genome to sequence to study low-level relationships, researchers are still unsure of how to weigh the effects of the slower mutation accumulation rate of cpDNA as compared to the longer coalescent time of nDNA.

Varying degrees of intraspecific cpDNA polymorphism have been encountered in numerous studies (Harris and Ingram 1991; Soltis et al. 1992). Most reports of conspecific

cpDNA variation represent instances of inter-, rather than intra-, population variation, but this trend could be the result of under-sampling individuals within each population (Soltis et al. 1992). More recently, quickly evolving chloroplast DNA sequence data has been applied in conjunction with an extensive, phylogeographic sampling scheme in order to infer evolutionary histories from the geographic distribution of cpDNA haplotypes (Shaw and Small 2005, Jakob and Blattner 2006, Naciri and Gaudeul 2007, Prentice et al. 2008).

The focus of this research was to screen five highly variable cpDNA gene regions to determine whether any of them display variation between individuals of *C. fremontii* S. Watson. The survey included both Midwestern and Southeastern populations and included individuals that, in an earlier preliminary study, displayed distinct nrDNA ITS alleles.

MATERIALS AND METHODS

***Taxon sampling*—**Modeled after the three species sampling scheme of Shaw et al. (2005), six accessions of *C. fremontii*, representing four eco-geographic regions, and one accession of *C. ochroleuca* were sampled in order to test the relative utility of three noncoding cpDNA markers, *trnG-trnG*, *trnD-trnT*, and *psbD-trnT*, in characterizing the genetic diversity within and among *C. fremontii* populations (Table 1). The relative utility of two other regions, *psbA-trnH* and *petA-psbJ*, was also tested but with different sampling schemes (Table 1). For *psbA-trnH*, nineteen *C. fremontii* accessions and one *C. ochroleuca* accession (n=20) were sampled, and for *petA-psbJ* 10 *C. fremontii* accessions were sampled (Table 1).

***DNA extraction, amplification and sequencing*—**I extracted DNA from leaves using the DNEasy Plant Mini Kit (Qiagen, Valencia, California, USA). I performed polymerase chain reaction (PCR) using Eppendorf (Westbury, New York, USA) Mastercycler gradient or Mastercycler personal thermal cyclers in 40 µl volumes with the following reaction components: 1 µL template DNA (~ 10-100 ng), 1x *ExTaq* buffer (PanVera / TaKaRa, Madison, Wisconsin, USA), 200 µmol/L each dNTP, 3.0 mmol/L MgCl2, 0.1 µmol/L each primer, and 1.25 units *ExTaq* (PanVera / TaKaRa). Reactions included bovine serum albumin at a final concentration of 0.2 µg/µL, which is known to improve amplification from difficult templates. Five sets of PCR and sequencing primers were used to amplify five target noncoding cpDNA regions (Table 2). The PCR protocol described next was preceded by template DNA denaturation at 80°C for 5

min and followed by a final extension step of 1 hour at 65°C. The PCR cycling conditions were 29 cycles of denaturation at 95°C for 15 sec., primer annealing at 50°C for 15 sec., followed by a ramp of 0.3°C/s to 65°C, and primer extension at 65°C for 5 min. I checked PCR products on 1% agarose gels before cleaning with ExoSAP-IT (USB, Cleveland, Ohio, USA). I performed DNA sequencing using forward and reverse primers of all 5 target regions using the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction kit, v. 3.1 (Perkin-Elmer/Applied Biosystems, Foster City, California, USA) and sequences were electrophoresed and detected on an ABI Prism 3100 automated sequencer (University of Tennessee Molecular Biology Resource Facility, Knoxville, Tennessee, USA). I used Sequencher 4.7 (Gene Codes, Ann Arbor, Michigan, USA) to edit the DNA strands and check any base calls at variable positions. All unique haplotypes sequences will be deposited in GenBank.

Alignments and phylogenetic analyses—I aligned DNA sequences by eye in MacClade v. 4.08 (Sinauer, Sunderland, Massachusetts, USA) and double-checked variable positions in the data matrix against the original chromatogram files to make sure that all base calls were true at all variable positions. In all cases but one, alignment of potentially informative positions was unambiguous; an approximately 350 bp portion of the *psbD-trnT* intergenic spacer region, located 600 bp after the start of the sequence marked by the *psbD* primer, contained two poly-A/T regions that interfered with sequencing. Therefore, this small region was excluded from the analysis as no confident alignment could be constructed. The number of informative characters was counted by eye in MacClade v. 4.08.

RESULTS

The length of each region sequenced and the number of variable sites displayed by each region sequenced are listed in Table 3. Polymorphic positions were minimal to nonexistent in all five regions examined. The region displaying the most single nucleotide polymorphisms was *trnD-trnT* (Table 3) with six mutations observed among four *C. fremontii* and one *C. ochroleuca* accessions. The *psbA-trnH* spacer displayed no single nucleotide polymorphisms; however, this was the only region for which *C. ochroleuca* was not included because PCR amplicons were generated prior to collection of subsection *Integrifoliae* taxa.

DISCUSSION

Unfortunately, results of the chloroplast screening were disappointing in that a minimal number of potentially informative characters was encountered across all five regions. Because each region screened displayed insignificant genetic variation among the *C. fremontii* accessions and *C. ochroleuca*, noncoding cpDNA gene regions are likely not going to provide sufficient information to tease apart population-level relationships among Midwestern and Southeastern populations of *C. fremontii*. Even though the chloroplast genome has been widely used in plant phylogeography to illuminate intraspecific variation (e.g., Morris et al. 2008), there are also studies where portions of this genome displayed too little genetic information to even separate species (Hardig et al. 2000). Plant genomes evolve at different rates; the mitochondrial genome evolves at the slowest rate, the chloroplast genome at a faster rate, and the nuclear genome at the fastest rate (Wolfe et al. 1987). Often the chloroplast genome does not accumulate mutations quickly enough for low-level taxonomic study, and even noncoding sequences and intergenic spacers can display insufficient variation at low taxonomic levels (Small et al. 1998), as evidenced by a phylogenetic study of *Clematis* in which one nrDNA region (ITS) was as informative as 5 non-coding cpDNA regions combined (Miikeda 2006). In light of the variable effectiveness of the chloroplast genome at low taxonomic levels, and the fact that the most variable regions of the cpDNA genome displayed minimal levels of variation among surveyed *Clematis* accessions, it appears noncoding cpDNA regions evolve at too slow a rate in *C. fremontii* to infer relationships among populations.

Because the cpDNA genome will not provide enough information to separate populations of *C. fremontii*, future work will incorporate nuclear DNA sequences. Specifically the next chapter assesses the variability of the common nuclear ribosomal DNA Internal Transcribed Spacer (ITS) region. Chloroplast DNA and nrDNA are popular molecular tools in phylogeographic investigations, and nrDNA is often used to offset the limitations of cpDNA (Small et al. 2004). Since the mitochondrial genome evolves too slowly and markers such as microsatellites, AFLPs, and allozymes are historically unordered, nuclear DNA is the only alternative to cpDNA that could provide important historical data regarding the genealogical relationships of allelic variants within these *C. fremontii* populations.

LIST OF REFERENCES

Avise, J.C. 2000. Phylogeography: The History and Formation of Species. Harvard University Press, Cambridge, MA.

Hardig, T. M., P. S. Soltis, and D. E. Soltis. 2000. Diversification of the North American shrub genus *Ceanothus* (Rhamnaceae): conflicting phylogenies from nuclear ribosomal DNA and chloroplast DNA. *American Journal of Botany* 87: 108-123.

Harris, S. A. and R. Ingram. 1991. Chloroplast DNA and Biosystematics: The Effects of Intraspecific Diversity and Plastid Transmission. Taxon 40(3): 393-412.

Jakob, S. S. and F. R. Blattner. 2006. A chloroplast genealogy of *Hordeum* (Poaceae): long-term persisting haplotypes, incomplete lineage sorting, regional extinction, and the consequences for phylogenetic inference. Molecular Biology and Evolution 23(8): 1602-1612.

Miikeda, O., K. Kita, T. Handa, T. Yukawa. 2006. Phylogenetic relationships of *Clematis* (Ranunculaceae) based on chloroplast and nuclear DNA sequencesBotanical Journal of the Linnean Society 152: 153–168.

Morris, A.B., S.M Ickert-Bond, D.B. Brunson, D.E. Soltis, and P.S. Soltis. 2008. Phylogeographical structure and temporal complexity in American sweetgum (Liquidambar styraciflua; Altingiaceae). Molecular Ecology 17(17): 3889-3900.

Naciri, Y. and M. Gaudeul. 2007. Phylogeography of the endangered *Eryngium alpinum* (Apiaceae) in the European Alps. Molecular Ecology 16: 2721-2733.

Prentice, H. C., J. U. Malm, and L. Hathaway. 2008. Chloroplast DNA variation in the European herb *Silene dioica* (red campion): postglacial migration and interspecific introgression. Plant Systematics and Evolution 272: 23-27.

Rosenberg, N. A. and M. Nordborg. 2002. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. Nature 3: 380-390.

Sang, T. 2002. Utility of low-copy nuclear gene sequences in plant phylogenetics. Cricital Reviews in Biochemistry and Molecular Biology. 37 (3): 121-147.

Schaal, B. A., J. F. Gaskin, and A. L. Caicedo. 2003. Phylogeography, haplotype trees, and invasive plant species. Journal of Heredity 94(3): 197-240.

Shaw, J., E. B. Lickey, J.T. Beck, S. B. Farmer, W. Liu, J. Miller, K. C. Siripun, C. T. Winder, E. E. Schilling, and R. Small. 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. American Journal of Botany 92(1): 144-166.

Shaw, J., E. B. Lickey, E. E. Schilling, and R. Small. 2007. The tortoise and the hare III: comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms. American Journal of Botany 94:275-288.

Shaw, J. and R.L. Small. 2005. Chloroplast DNA phylogeny and phylogeography of the North American plums (*Prunus* subgenus *Prunus* section *Prunocerasus*, Rosaceae). American Journal of Botany 92(12): 2011-2030.

Small, R. L., J. A. Ryburn, R. C. Cronn, T. Seelanan, and J. F. Wendel. 1998. The tortoise and the hare: choosing between noncoding plastome and nuclear *Adh* sequences for phylogeny reconstruction in a recently diverged plant group American Journal of Botany 85: 1301-1315.

Small, R. L., R. C. Cronn Richard C., and J. F. Wendel. 2004. L. A. S. JOHNSON REVIEW No. 2. Use of nuclear genes for phylogeny reconstruction in plants. Australian Systematic Botany 17: 145–170.

Soltis, D.E., P.S. Soltis, and B.M. Milligan. 1992. Intraspecific chloroplast DNA variation: systematic and phylogenetic implications. Pp. 117-150 in P.S. Soltis, D.E. Soltis, and J.J. Doyle, eds. Molecular systematics of plants. Chapman and Hall, New York.

Tate, J. A. and B. B. Simpson. 2003. Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. Systematic Botany 28: 723-737.

Wolfe, K. H., W. H. Li, and P. M. Sharp. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc. Nat. Acad. Sci. 84: 9054-9058.

Table 1. Taxa used in this investigation, source and voucher numbers, total DNA numbers, and cpDNA regions sequenced. Successful sequence generated for each cpDNA region is represented with an asterisk. Each cpDNA region is named according to coding sequences flanking the region.

					cpDNA				
					trnG-trnG	*trnD-trnT*	*psbD-trnT*	*psbA-trnH*	*petA-psbJ*
Species	Collector; Number; Year	Herb.	State and County	Total DNA #					
C. fremontii S. Watson	C. Montgomery; CMo1; 2007	UCHT	USA; TN; Hamilton						
				563				*	*
				564				*	*
				565				*	*
				566				*	
				567				*	
				568	*	*	*	*	*
				569				*	
				570	*	*	*	*	
				571				*	
				572				*	
				573				*	
				574				*	
				575				*	
				576				*	
				577					
				578				*	
				579					
				580				*	

Table 1. Taxa used in this investigation, source and voucher numbers, total DNA numbers, and cpDNA regions sequenced. Successful sequence generated for each cpDNA region is represented with an asterisk. Each cpDNA region is named according to coding sequences flanking the region.

					cpDNA				
					trnG-trnG	*trnD-trnT*	*psbD-trnT*	*psbA-trnH*	*petA-psbJ*
Species	Collector; Number; Year	Herb.	State and County	Total DNA #					
C. fremontii S. Watson	C. Montgomery; CMo1; 2007	UCHT	USA; TN; Hamilton						
				581				*	
				582				*	
				583				*	
C. fremontii S. Watson			USA; GA; Floyd						
				588					*
				590					*
				596	*	*	*		
C. fremontii S. Watson	C. Montgomery; CMo2; 2008	UCHT	USA; MO; Jefferson						
				602					*
C. fremontii S. Watson	C. Montgomery; CMo5; 2008	UCHT	USA; MO; Washington						
				618					*
C. fremontii S. Watson	C. Montgomery; CMo6; 2008	UCHT	USA; MO; Ozark						
				622	*	*	*		
C. fremontii S. Watson	C. Montgomery; CMo10; 2008	UCHT	USA; KS; Phillips						
				644	*	*	*		*

Table 1. Taxa used in this investigation, source and voucher numbers, total DNA numbers, and cpDNA regions sequenced. Successful sequence generated for each cpDNA region is represented with an asterisk. Each cpDNA region is named according to coding sequences flanking the region.

					cpDNA				
					trnG-trnG	*trnD-trnT*	*psbD-trnT*	*psbA-trnH*	*petA-psbJ*
Species	Collector; Number; Year	Herb.	State and County	Total DNA #					
C. fremontii S. Watson	C. Montgomery; CMo7; 2008	UCHT	USA; KS; Jewell						
				650	*	*	*		*
C. ochroleuca Aiton	C. Montgomery; CMo18; 2008	UCHT	USA; GA; Stephens						
				651	*	*	*	*	

Table 2. Chloroplast DNA regions used in this investigation and the corresponding forward and reverse primer names and sequences. For each region forward primers are listed before reverse primers.

cpDNA Region	Primer name and sequence (5'-3')
psbA-trnH	**trnH$^{(GUG)}$:** CGC GCA TGG TGG ATT CAC AAT CC (Tate and Simpson, 2003)
	psbA: GTT ATG CAT GAA CGT AAT GCT C (Sang et al., 1997)
psbD-trnT	**psbD:** CTC CGT ARC CAG TCA TCC ATA (Shaw et al., 2007)
	trnT$^{(GGU)}$-R: CCC TTT TAA CTC AGT GGT AG (Shaw et al., 2007)
psbJ-petA	**psbJ:** ATA GGT ACT GTA RCY GGT ATT (Shaw et al., 2007)
	petA: AAC ART TYG ARA AGG TTC AAT T (Shaw et al., 2007)
trnD^{GUC}-trnT^{GGU}	**trnD$^{(GUC)}$F:** ACC AAT TGA ACT ACA ATC CC (Demesure et al.,1995)
	trnT$^{(GGU)}$: CTA CCA CTG AGT TAA AAG GG (Demesure et al.,1995)
trnG^{UUC}-trnG^{UUC}	**trnS$^{(GCU)}$:** AGA TAG GGA TTC GAA CCC TCG GT (Shaw et al., 2005)
	5'trnG2G$^{(GCG)}$: GGT ATA GTT TAG TGG TAA AA (Shaw et al., 2005)
	3'trnG$^{(UUC)}$: GTA GCG GGA ATC GAA CCC GCA TC (Shaw et al., 2005)

Table 3. Approximate length for each region sequenced in base pairs and number of variable sites. Because of two poly A/T regions that interfered with sequencing, for the purpose of sequence alignment, approximately 350 bp of continuous sequence was removed from all sequences of the *psbD-trnT* region; the length reported for the *psbD-trnT* region includes the ~350 bp portion of ambiguous sequence. The number of variable characters in aligned sequence is reported for each region.

		Number of		
Region	Approximate Length (bp)	intraspecific variable sites	interspecific variable sites	total variable sites
trnG^{UUC}-*trnG*^{UUC}	712	1	0	1
trnD^{GUC}-*trnT*^{GGU}	1168	2	4	6
psbD-trnT	1660	2	1	3
psbA-trnH	358	1	1	2
psbJ-petA	493	0	NA	0

Part 3. Incomplete concerted evolution in the non-hybrid diploid *Clematis fremontii*

INTRODUCTION

The nuclear ribosomal DNA (nrDNA) ITS region is among the highly repetitive class of DNA, with copy number in eukaryotes ranging from several thousands to millions per diploid genome (Long and Dawid 1980). In eukaryotes, copies of the nrDNA gene, consisting of a transcription unit and non-transcribed spacer, are tandemly repeated throughout the nuclear genome within one or a few chromosomal loci (Long and Dawid 1980). The internal transcribed spacers (ITS-1 and ITS-2) separate the 18S, 5.8S, and 28S exons that code for ribosomal subunits. Located upstream of the 18S gene is the external transcribed spacer (ETS). Each adjacent nrDNA transcription unit consists of these three exons (18S, 5.8S, 28S) and three transcribed spacers (ITS-1, ITS-2, ETS) and is separated by a non-transcribed spacer (NTS) (Figure 1).

Despite the high copy number, the nrDNA gene is putatively conserved within individuals because repeats are subject to the processes of molecular drive and subsequent concerted evolution (Elder and Turner 1995). Unequal crossing over and gene conversion (transposition and reamplification of a DNA sequence) are the molecular drive mechanisms thought to govern concerted evolution, i.e. the nonindependent evolution of repetitive DNA sequences (Dover, 1982). In theory, concerted evolution leads to individuals with homogeneous sequence in high copy number, making the transcribed spacer units, which are less functionally constrained than the nrDNA exons, ideal molecular markers for low level phylogenetic study.

Internal transcribed spacer sequence data has a multitude of putatively advantageous properties for phylogenetic inference in plants. Four characteristics of ITS provide the allure expeditious laboratory procedure. (1) High amplicon yield and (2) single-primer sequencing are facilitated by high nrDNA copy number. (3) The relatively short length, 500-700 bp in angiosperms (Baldwin et al. 1995), means that clean, bidirectional sequences of the region may be easier to obtain, relative to longer and low-copy nuclear genes. (4) There exists a set of universal primers for the ITS region that are effective across most plant phyla (White et al. 1990). Three other beneficial properties facilitate low level phylogenetic study. Biparental inheritance of the nuclear genome provides more lineage information per individual than the maternally inherited chloroplast genome (Alvarez and Wendel 2003). ITS-1 and ITS-2 are presumably under mild functional constraint which enables generous and essentially neutral evolution of nucleotide sites (Alvarez and Wendel 2003). Lastly, concerted evolution of

ribosomal cistrons purportedly offers the advantages of intragenomic homogeneity and intergenomic variability, i.e. a lack of confounding intraindividual variation and the presence of potentially phylogenetically informative interindividual variation.

Use of the ITS region dates back to the infancy of molecular phylogenetic research (Alvarez and Wendel 2003), and the presumed benefits of ITS sequence data likely contribute substantially to its extensive application in plant phylogenetics (Bailey et al. 2003). That ITS has been one of the most important gene regions for relational studies in plants is exemplified by a Web of Science citation search of the paper that characterized this gene region (Baldwin et al. 1995), which showed that it has been cited 900 times. A search on NCBI-GenBank for "internal transcribed spacer region" resulted in nearly 610,000 nucleotide sequences. However commonplace the application of this region may be, the evolutionary mechanisms acting on nrDNA arrays and the subsequent effects on phylogenetic relationship inference remain obscure, which has led to several cries of caution regarding the use of ITS for phylogenetic studies. Intraindividual nrDNA polymorphism is primarily responsible for these warnings, as divergent intraindividual gene copies can be caused by incomplete lineage sorting of ancestral homologues, recombination between copies, pseudogenes or divergent and unconcerted paralogous copies, or almost random fixation of any of the thousands of copies (Buckler et al., 1997; Bailey et al., 2003). Most of the potential phylogenetic complications associated with the use of ITS sequences can be linked to the molecular architecture of the nrDNA cistron (Alvarez and Wendel 2003).

Counter to early blanket claims that concerted evolution leads to homogenization of nrDNA repeats in individuals (e.g., Baldwin et al. 1995; Elder and Turner 1995; Hillis and Dixon 1991, Ainouche and Bayer 1997), reports of extensive intra-genomic polymorphism are on the rise. Several accounts of intra-genomic polymorphism in hybrid diploids (e.g. Noyes 2006; Feliner et al. 2004; Weeks and Simpson 2004; Wichman et al. 2002; Rieseberg et al. 1991; Sang et al. 1995; Quijada et al. 1997; Kita and Ito 2000) and polyploids (e.g. Devos et al. 2005; Wendel et al. 1995; Rauscher et al. 2004; Koch et al. 2003) have been published. These studies emphasized the idea that concertion is usually incomplete in relatively recent hybrid or polyploid taxa. More specifically, these studies compared ITS alleles of the subject taxa with those of putative or confirmed parental taxa to reveal species history and or investigate the signature of concerted evolution within hybrid species.

Few instances of intragenomic polymorphism have been observed in putative non-hybrid diploids (Hughes et al., 2003; Harpke and Peterson, 2006; King and Roalson, 2008; Saini et al., 2008). By cloning individuals displaying ITS sequence polymorphism, Hughes et al. (2003) found four instances of intraindividual polymorphism within three individuals, representing three species of *Desmanthus* Willd. and one individual of *Leucaena pulverulenta* Benth. Although the ITS copies from the *L. pulverulenta* accession were monophyletic with other accessions of the same species on an ITS gene tree, the ITS copies from the three *Desmanthus* species were polyphyletic. Despite the polyphyly of the three *Desmanthus* species, Hughes et al. (2003) generated a well-supported ITS phylogeny of the informal, *Dichrostachys* and *Leucaena* groups of the Mimosoideae. Using cDNA, PCR amplified sequences, and clones, Harpke and Peterson (2006) found a high degree of intraindividual polymorphism within the ITS region of 21 species (wild and cultivated), representing various intrageneric levels of *Mammillaria* Haw. (Cactaceae). In multiple instances, genomic DNA ITS sequences from the same individual were polyphyletic, and in one instance intraindividual variation was as high as 32%. Harpke and Peterson (2006) concluded extensive intraindividual polymorphism in *Mammillaria* represented non-concerted evolution of nrDNA, and they cautioned against the use of ITS in further phylogenetic investigation of the genus. Through extensive cloning, King and Roalson (2008) identified high levels of intraindividual polymorphism of both ITS and ETS sequences in 87 individuals, representing 73 species of subgenus *Vignea* (P.Beauv.) A.E. Kozhevn. of *Carex* L. (Cyperaceae). Based on the high number of paralogous nrDNA copies within subg. *Vignea* as well as polyphyly of ITS sequences from *C. cephalophora* Muhl. ex Willd. and paraphyly of ITS sequences from *C. gibba* Wahlenb. with a separate species, King and Roalson (2008) concluded that incomplete concerted evolution of nrDNA complicates phylogenetic inference within that subgenus. Saini and others (2008) found intraindividual polymorphism among wild and cultivated accessions of *Vigna radiata* (L.) R. Wilczek and concluded that intragenomic polymorphism was a result of intraspecies hybridization between wild and cultivated types of *V. radiata*. In all of the aforementioned studies, careful attention was paid to pseudogene identification, with the exception of Hughes et al. (2003); however Hughes did coauthor *Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes* (Bailey et al. 2003).

Pseudogenes are non-functional nrDNA copies, which are able to accumulate mutations at a quicker rate than functional ribosomal cistrons. In light of the molecular structure and

evolution of the ribosomal subunits, assuming orthology in ITS sequence data is dubious (Alvarez and Wendel 2003, Small et al. 2004, Feliner and Rosselló 2007). In order to curtail the risk associated with ITS generated phylogenies, many have extended their research to include pseudogene analyses (e.g. Bailey et al., 2003; Harpke and Peterson, 2006, 2007; King and Roalson, 2008; Saini et al., 2008; Wei et al., 2003). Several methods have been utilized for the detection and elimination of pseudogenes, which include assessing relative concentrations of G-C content (Razafimandimbison et al. 2004), free energy calculation (Buckler et al. 1997), analysis of secondary rRNA structure (Harpke and Peterson 2006), analysis of relative rates of nucleotide substitution between introns and exons (Bailey et al. 2003), and identification of conserved Viridiplantae 5.8S motifs (Harpke and Peterson 2008). Irrespective of the phylogenetic complications associated with nrDNA, ITS sequences can be a useful tool for uncovering instances of incomplete lineage sorting, introgression, hybridization, and reticulation (King and Roalson 2008; Okuyama et al. 2005; Koch et al. 2003; Morgan 2003) and despite the myriad complications this gene region is still among the most widely used today. This is exemplified by the title of a paper by Feliner and Rosselló (2007), entitled "Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants," in which the authors champion the use of ITS and the insightful results it can render, but suggest that it be used with caution in light of the complications associated with the region.

Comprised of about 280 species, *Clematis* L., has a worldwide distribution and is among the largest genera in Ranunculaceae (Yang and Moore 1999). Ranunculaceae is considered by most botanists to be among the most basal groups of eudicots (APG II, 2003). Within Ranunculaceae, *Clematis* species have single-seeded achenes that may or may not be covered with hairs. Chromosomal counts of 42 different *Clematis* species revealed that *Clematis* is a diploid genus with eight pairs of chromosomes (2n=16) (Gregory 1941). Polyploidy is rare in *Clematis* but has been documented in three hybrid cultivars (Erickson 1943); it should be noted that polyploidy has not been observed in wild uncultivated *Clematis*. While *Clematis* is used widely in horticulture and ranks among the most popular ornamental plants, many *Clematis* species, such as North American members of Erickson's (1943) subsection *Integrifoliae* (*C. ochroleuca* Aiton, *C. albicoma* Wherry, *C. viticaulis* Steele, *C. coactilis* (Fernald) Keener, and *C. fremonii* S. Watson) of subg. *Viorna* (Rchb.) Tamura, are rarely used in horticulture (Grey-Wilson 2000).

A major focus of this investigation is on *C. fremontii*, a long-lived, diploid (2n=16) perennial with a distribution that is limited to isolated populations on dolomitic or limestone substrates of calcareous glades and prairies in Missouri and Kansas (Erickson 1943, Gregory 1941). Isolation of populations is likely absolute with virtually no interbreeding or seed dispersal, as populations are separated by several miles, animal dispersal is unlikely and seeds lack the feathery plumes necessary to travel great distances via wind dispersal. *Clematis fremontii* is of primary interest because two disjunct populations were recently discovered in Hamilton County, Tennessee and Floyd County, Georgia.

At the outset of this course of study the main objective was to use the ITS region to assess the origin of the two disjunct southeastern *C. fremontii* populations as well as to investigate the phylogeography of *C. fremontii* in relation to other taxa in Erickson's (1943) subsection *Integrifoliae* species within the eastern United States: *C. coactilis*, *C. albicoma*, *C. viticaulis*, and *C. ochroleuca.* To that end, ITS sequences of the southeastern populations of *C. fremontii* were compared to naturally occurring midwestern populations and other wild subsection *Integrifoliae* taxa. Although the original intent of this research was to assess the origin of the two disjunct southeastern *C. fremontii* populations, an unusually high frequency of intraindividual polymorphism of the ITS region prompted this investigation into the extent of these alleles throughout this species, section *Integrifoliae*, subgenus *Viorna*, and the rest of the genus. The focus of the research presented here is an exploration, through extensive cloning and pseudogene analysis, into the occurrence of intragenomic polymorphism in nrDNA of *C. fremontii* and related taxa within subsection *Integrifoliae*.

MATERIALS AND METHODS

***Taxon sampling*—**A total of 153 accessions were collected, representing nine species of subg. *Viorna* (Table 1). One hundred and thirty of those individuals were *C. fremontii*, including 22 individuals from the single population in Tennessee, 13 individuals from the single population in Georgia. Ten individuals per population from four populations and five individuals from one population were sampled from Missouri, and five individuals per population from five populations were sampled from Kansas. One population was sampled per species for the remaining subsect. *Integrifoliae* taxa, with a sampling scheme of 5 individuals per population of: *C. coactilis, C. viticaulis, C. albicoma,* and *C. ochroleuca* and one individual per population

from subsect. *Viorna* taxa: *C. glaucophylla*, *C. viorna*, *C. addisonii*. All plants were collected from wild populations (Figure 2).

To maintain order among the cloned accessions (individuals), populations, and species the following naming system was adopted where the first letter designates the species (F=*C. fremontii*, O=*C. ochroleuca*, C=*C. coactilis*, A=*C. albicoma*, V=*C. viticaulis*, D=*C. addisonii*, G=*C. glaucophylla*, and R=*C. viorna*), which is followed by the state abbreviation (GA, KS, MO, TN, VA, NC, AL), population number (refer to Fig. 2), a period, accession number, and a letter identifying the clone (e.g., FTN1.570A) .

Sixty-two individuals from the153 total accessions of *C. fremontii*, representing all 10 of the sampled populations, plus 23 individuals from one population each of the remaining eight species sampled within subg. *Viorna* were selected for cloning. The 62 *C. fremontii* individuals were selected based on polymorphic characters observed in the direct sequencing of PCR amplicons of the (ITS1-5.8S-ITS2) region, since the original intent of selection was to determine the alleles present in apparent heterozygotes. Five individuals per population were chosen for cloning from ten midwestern populations and from the Georgia population, and seven individuals were selected for cloning from the Tennessee population. In order to include ample sister taxa in the clone dataset, the 23 non-*C. fremontii* subsect. *Integrifoliae* species individuals were cloned irrespective of ambiguous characters. Of the 85 cloned individuals, a minimum of four clones per individual were sequenced. Four individuals of *C. fremontii*, one from each of four main population areas, FTN1.570, FGA1.590, FMO1.604, and FKS2.633, were chosen for a more thorough sampling of clones and in each of these individuals 20 clones were sequenced.

DNA extraction, amplification, and sequencing—DNA was extracted from leaves using the DNeasy Plant Mini Kit (Qiagen, Valencia, California, USA). The polymerase chain reaction (PCR) was performed using Eppendorf Mastercycler gradient cycler in 25 µl volumes with the following reaction components: 1 µL template DNA (~ 10-100 ng), 1X ExTaq buffer (PanVera / TaKaRa, Madison, Wisconsin, USA), 200 mmol/L each dNTP, 3.0 mmol/L MgCl2, 0.1 µmol/L each primer, and 1.25 units ExTaq (PanVera / TaKaRa). Reactions included bovine serum albumin at a final concentration of 0.2 mg/mL, which is known to improve amplification from difficult templates. PCR and sequencing primers are listed in Table 2. The PCR protocol described below was preceded by template DNA denaturation at 80°C for 5 minutes and

followed by a final extension step of 5 min at 65°C. The PCR cycling conditions were 30 cycles of: denaturation at 94°C for 1 min, primer annealing at 53°C for 1 min, and primer extension at 72°C for 1min. PCR products were checked on 1% agarose gels before being cleaned with ExoSAP-IT (USB, Cleveland, Ohio, USA). DNA sequencing was performed using both primers. All DNA sequencing was performed with the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit, v. 3.1 (Perkin-Elmer / Applied Biosystems, Foster City, California, USA) and electrophoresed and detected on an ABI Prism 3100 automated sequencer (University of Tennessee Molecular Biology Resource Facility). Sequencher 3.1.1 (Gene Codes) was used to edit the DNA strands. Sequences of unique haplotypes will be deposited in GenBank.

***Cloning*—**Amplification products were visualized by agarose gel electrophoresis before being cleaned with the QIAquick® PCR Purification Kit (Quiagen). Purified amplification products were ligated into the pGEM-T Easy vector and transformed into JM109 competent *E. coli* cells (Promega, Madison, Wisconsin, USA) according to the manufacturer's instructions. Ten of the resulting colonies were screened for plasmids via PCR using the vector specific T7 and SP6 primers; 30 colonies were screed for the individuals chosen for exhaustive cloning (Table 3). This PCR amplification protocol was preceded by template DNA denaturation at 80°C for 5 minutes and followed by a final extension step of 10 min at 65°C. The PCR cycling conditions for amplifying the cloned fragments were 25-30 cycles of: denaturation at 94°C for 1 min, primer annealing at 50°C for 1 min, and primer extension at 72°C for 2 min. From the 10 PCR amplified clones of each accession, a minimum of four that yielded bands of the appropriate size (~950 bp) were chosen for sequencing in an attempt to obtain both alleles in the heterozygotes and a minimum number of four cloned (ITS1-5.8S-ITS2) amplicons. In the instance of exhaustive cloning, appropriately sized bands were chosen for sequencing until a minimum of 20 clones were successfully sequenced. In some cases, shorter bands (~680 bp and ~520 bp) were viewed on the gel. These bands were checked against a BLAST search to determine whether or not they were ITS sequences. The PCR amplicons were then cleaned using ExoSAP-IT (USB, Cleveland, Ohio, USA) and sequenced with a single primer (T7). All DNA sequencing was performed commercially by Macrogen, Inc (Seoul, South Korea) or was performed with the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit, v. 3.1 (Perkin-Elmer / Applied

Biosystems, Foster City, California, USA) and electrophoresed and detected on an ABI Prism 3100 automated sequencer (University of Tennessee Molecular Biology Resource Facility). Sequencher 3.1.1 (Gene Codes) was used to edit the DNA strands. All sequences will be deposited in GenBank.

Alignment and Phylogenetic analyses—In addition to the sequences generated for this study, all available ITS sequences of *Clematis* species were downloaded from NCBI-GenBank and included within the aligned dataset (Table 4); the majority of GenBank accessions were from Miikeda et al. (2006). The tree was rooted with *C. nobilis* Nakai because this was one of the species suggested to be basal in the total evidence tree of Mikeda et al. (2006). DNA sequences were aligned by eye in MacClade v. 4.0 (Sinauer, Sunderland, Massachusetts, USA). Variable positions in the data matrix were double checked against the original chromatogram files to verify that all base calls were true at all variable positions. In all cases but two, alignment of potentially informative positions was unambiguous. In each case two bp were deleted due to a poly C run located 442bp and 489bp from the 5' end of the sequences. Also, within the outgroup species sequences downloaded from GenBank there was a 24 bp portion, located in the ITS1 region 100 bp upstream from the 5' end of the 5.8S exon, that was deleted among these species due to ambiguous alignment. Indels were coded as binary characters.

An analysis using all of the ITS sequences generated by directly sequencing PCR products plus the available GenBank ITS sequences was performed. This analysis was conducted using the optimality criterion of maximum parsimony. Searches for most-parsimonious trees were executed in PAUP* v. 4.0 b10 (Swofford, 2002) by a heuristic search limited to 10000000 rearrangements with tree bisection-reconnection (TBR) branch swapping and the "collapse zero-length branches" option in effect.

The analysis of cloned sequences was performed using the optimality criterion of maximum parsimony. Searches for most-parsimonious trees were executed in PAUP* v. 4.0 b10 (Swofford, 2002) by a heuristic search limited to 10000000 rearrangements with tree bisection-reconnection (TBR) branch swapping and the "collapse zero-length branches" option in effect.

Pseudogene analysis—In the study, *5.8S Motifs for the Identification of Pseudogenic ITS Regions*, Harpke and Peterson (2008) found the Magnoliophyta motif M1 as effective in

identifying pseudogenes in two datasets of distantly related genera, 70 5.8S sequences of *Quercus* (Fagaceae) and 65 5.8S sequences of *Mammillaria* (Cactaceae), as three other commonly used, but much more complex methods of pseudogene identification, (1) 5.8S secondary structure reconstructions, (2) the relative rate test (Takezaki et al. 1995), and (3) tree-based bootstrap hypothesis testing (Bailey et al. 2003). Therefore, in this study the 5.8S sequences of all clones (Table 3) was searched for presence of the 16 bp Magnoliophyta motif M1 (CGATGAAGAACGTAGC) (Harpke and Peterson 2008), and any clone lacking this motif was determined a pseudogene.

Intraindividual polymorphism—To characterize haplotypes and identify instances of intraindividual polymorphism, sequences of the clone dataset (Table 3) were collapsed into haplotypes using COLLAPSE version 1.2 using the –g –m command (Posada 2005).

RESULTS

Nonspecific primer binding—The direct PCR of genomic DNA for all 82 accessions revealed a single dominant band of approximately 750 bp. Upon cloning, two smaller bands (~680 bp and ~520 bp) were revealed in some individuals. A BLAST search of cloned sequences from these small bands identified them as the chloroplast gene *rrn23* which codes for 23S rRNA and the mitochondrial gene *rrn26* which codes for 26S rRNA, indicating the primers ABITS101 and ABITS102 can bind to 3 different ribosomal RNA genes in these three separate plant genomes. Identification of the smaller bands as non-nrDNA ITS supports the methodology of only selecting appropriately sized bands to sequence in the cloning effort.

Phylogenetic analysis (Direct PCR)—The parsimony analysis of ITS1-5.8S-ITS2 sequences of the 711 bp alignment of 82 sequences from accessions of *Clematis* taxa (62 *C. fremontii*, five *C. albicoma*, five *C. coactilis*, five *C. viticaulis*, three *C. ochroleuca*, one *C. viorna*, and one *C. addisonii*) and ITS sequences from 38 *Clematis* taxa from GenBank (Table 4) yielded the majority rule consensus phylogenetic tree in Figure 3. The data set contained 711 total characters; 125 characters were variable and 61 were parsimony informative. There were 15 variable characters among the subg. *Viorna* species; however only four were parsimony informative and support clades B, C, D, and E. At the subgeneric rank, Miikeda et al. (2006)

reported that all four subgenera (*Flammula*, *Viorna*, *Clematis*, and *Campanella*) were para/polyphyletic, and our results, based on largely the same data set, agree with that assessment. Clade A includes all North American subg. *Viorna* taxa, although the Asian temperate *C. fusca* of subg. *Viorna* is unresolved from the rest of the species of subg. *Viorna*. The data presented here show *C. texensis*, a Texas endemic, and *C. viorna* to be deeply nested within the species *C. fremontii*. All of the *C. fremontii* accessions, with the exception of FTN1.570 group together in Clade F. Within Clade F is Clade D, which contains both *C. texensis* and *C. fremontii* along with accessions of *C. fremontii* from all four major population areas (TN, GA, KS, and MO).

Intraindividual polymorphism identified through cloning—The clone data set was collapsed from 473 sequences to 333 unique haplotypes (Table 3). The number of haplotypes that are unique to one or more clones of a single accession is 321. Twenty-one haplotypes had a frequency greater than one, however only 12 haplotypes represented clones from more than one accession (Table 5). In ascending order according to number of accessions per haplotype, haplotype numbers 155, 210, 171, and 251 had a frequency of two accession, haplotype number 200 had a frequency of three, haplotype numbers 16 and 211 had a frequency of four, haplotype number 4 had a frequency of seven, haplotype number 5 had a frequency of 11, haplotype number 47 had a frequency of 13, haplotype number 24 had a frequency of 20, and haplotype number 2 had a frequency of 23.

Half of the 12 haplotypes that included multiple accessions also included multiple species. Haplotype # 155 included clones from one accession of *C. fremontii* (population FMO4) and one accession of *C. albicoma* subsect. *Integrifoliae.* Haplotype # 210 included clones from two accessions from two different species, *C. glaucophylla* subsect. *Viorna* and *C. ochroleuca* subsect. *Integrifoliae*. Haplotype # 211 included clones from one accession of *C. viticaulis* subsect. *Integrifoliae* and 3 accessions of *C. albicoma*. Haplotype #5 included clones from 10 accessions of *C. fremontii* and 1 accession of *C. albicoma*. Haplotype #47 included seven accessions of *C. fremontii*, three accessions of *C. coactilis* subsect. *Integrifoliae*, one accession of *C. viticaulis*, one accession of *C. ochroleuca*, and one accession of *C. viorna* subsect. *Viorna*.

The remaining six haplotypes are shared by more than one accession of the same species. One haplotype is unique to three accessions of *C. ochroleuca*, and the other five are unique to *C.*

fremontii. Haplotype #2 is unique to *C. fremontii* and includes the highest number of accessions out of all 333 haplotypes. Haplotype #2 contains accessions from every population of *C. fremontii* sampled; however the accession FTN1.570 from the Tennessee population was only identified as possessing this haplotype after the more thorough cloning of this individual (see below). Haplotype # 4 represents seven accessions from FGA1 (2 accessions), FMO3, FMO5, FKS1, and FKS4 (2 accessions). Haplotype #16 represents four accessions from FTN1, FGA1, FMO1, and FKS5. Haplotype #200 represents three individuals from FTN1. Haplotype #171 represents 2 accessions from FGA1 and FMO5.

Phylogenetic analysis of the cloned data set—The ML analysis of ITS1-5.8S-ITS2 sequences of the 711 bp alignment of 474 clones of ITS sequences representing 85 *Clematis* accessions (62 *C. fremontii*, five *C. albicoma*, five *C. coactilis*, five *C. viticaulis*, five *C. ochroleuca*, one *C. viorna*, one *C. glaucophylla*, and one *C. addisonii*) and ITS sequences from 38 *Clematis* taxa from GenBank (Table 4) yielded the phylogenetic tree in Figure 4. In this data set there were 740 total characters and 471 were variable. Among the variable characters over half of them, 269, were parsimoniously informative.

Results regarding the relationships of the species outside of subg. *Viorna* are in line with those reported by Miikeda et al. (2006) in that all subgenera are para/polyphyletic, and only North American subg. *Viorna* taxa form a monophyletic clade (Clade V) (with the exception of *C. fusca*). Clades V-V'''' is a grade containing individuals from multiple different species from two subsections, *Viorna* and *Integrifoliae*, neither of which is monophyletic. At the species rank, no monophyletic clades are present, as all of the clones from section *Integrifoliae* species resolve to different clades. Clades C, D, E, O, R, T and X are the larger clades containing only clones from *C. fremontii*. Clades C and D contain individuals from every sampled population of this species. Clade R includes only individuals of *C. fremontii* from the Georgia and Tennessee populations, which are the only two populations known to occur east of the Mississippi River. Individuals of *C. fremontii* are also included in several clades that contain multiple species. There are many clades in which clones from *C. fremontii* are grouped with clones from the other species in section *Integrifoliae*; for example, clades B, K, N, Q, Y, S, Y, and U. Clade Y includes both western subg. *Viorna* taxa, *C. fremontii* and *C. texensis*, and one individual of *C.*

ochroleuca collected in Georgia. All *C. fremontii* clones included within clade Y are from individuals from three Kansas populations.

Of the remaining subg. *Viorna* clones, only *C. ochroleuca* and *C. coactilis* formed monotypic clades at the species level where several (but not all) of the clones segregated into a clade. *Clematis ochroleuca* formed two clades (M and P) unique to the species; however these clades are in different parts of the tree. *Clematis coactilis* formed one clade (H) unique to the species. Like *C. fremontii*, all other clones of subg. *Viorna* taxa were included in numerous clades that contained multiples species. Clade L contained clones from all subsect. *Integrifoliae* species with an eastern distribution, *C. albicoma*, *C. coactilis*, *C. viticaulis*, and *C. ochroleuca*, as well as one clone from *C. addisonii*, an eastern subsect. *Viorna* species. Within L, one clone from *C. addisonii* and one clone from *C. viticaulis* formed a tertiary clade. Within Q, one clone from the same *C. addisonii* individual was grouped with 30 clones from *C. fremontii*. A third clone from the same *C. addisonii* individual was included in Clade Z with two individuals of *C. albicoma* and two individuals of *C. viticaulis*.

***Pseudogene identification*—**Using the M1 conserved 5.8S motif (Harpke and Peterson 2008), 18 clones were identified as pseudogenes. Fourteen pseudogenes were identified in 11 individuals of *C. fremontii*. Three pseudogenes were observed in the FTN1 population, one in FMO2, FKS1, and FKS3, two in FMO3, FMO4, FMO5, and FKS2; no pseudogenes were observed in FGA1, FMO1, FKS4, and FKS5. Three of the 14 clones identified as pseudogenes were captured through the more thorough cloning of the individuals FTN1.570 (two pseudogenes) and FKS2.633 (one pseudogene). One pseudogene was identified in an accession of *C. albicoma* and three pseudogenes were identified in a single accession of *C. coactilis*.

Putative pseudogenes are marked across the phylogeny with red dots (Fig. 4) and six of the 14 putative pseudogenes clustered together into Clade S, which includes six putative pseudogenes from five *C. fremontii* individuals and four putative functional ITS copies from two individuals of *C. fremontii* from separate Missouri populations and from one individual of *C. albicoma* and one individual of *C. viticaulis*. Within Clade S, the pseudogenes that form a single clade are marked with a large red dot (Fig. 4).

***Exhaustive cloning*—**In an attempt to survey how many alleles might be present within an individual a more thorough sampling of clones was initiated in four individuals, one from each major area of collection (GA- FGA1.590, KS- FKS2.633, MO- FMO1.604, TN - FTN1.570). Out of 96 clones from these four individuals, 67 had a unique haplotype, and 29 displayed one of nine shared haplotypes. The individual from Kansas (FKS2.633) displayed the highest number of clones with unique haplotypes (19) and the lowest number of shared haplotypes; three clones displayed haplotype #2 and two clones displayed haplotype #47. Individuals FTN1.570 and FMO1.604 both displayed four shared haplotypes, and individual FGA1.590 displayed five unique haplotypes. Individuals FTN1.570 had 14 unique clones (haplotypes); FGA1.590 had 13 unique clones, and FMO1.604 displayed 11 unique clones.

DISCUSSION

The occurance of extensive intraindividual polymorphism and widespread polyphyly within lineages in this study is indicative of incomplete concerted evolution, incomplete lineage sorting, and deep paralogy within North American sub. *Viorna* at the intraindividual and interspecific levels. Unfortunately, the dataset presented here is not informative enough to attempt to separate these three phenomena.

According to Elder and Turner (1995), natural populations restricted by limited gene flow and reproductive isolation should undergo concerted evolution to completion. The midwestern *C. fremontii* accessions are putatively diploid, naturally occurring reproductively isolated populations, yet in these accessions, the high frequency of more than two unique copies of the ITS region per individual indicates broadscale incomplete concerted evolution of nrDNA repeats. The extensive intraindividual polymorphism identified in all taxa of North American subg. *Viorna* indicates incomplete concerted evolution unequivocally characterizes subsect. *Integrifoliae* and likely extends to all of subg. *Viorna*. However, confirmation of incomplete concerted evolution within the entire subgenus would require more sampling of taxa outside of subsect. *Integrifoliae*.

The intraindividual polymorphism is so vast that 333 different ITS copies were found in 473 clones of only eight closely related species (85 accessions). In a comparison of mutant ratios, using the Kunkel method, the proofreading Taq polymerase *ExTaq* (PanVera / TaKaRa), used in this study, shows a four-fold higher fidelity than *Taq* (PanVera / TaKaRa), which has an

error rate of 8.7×10^{-6} (ExTaq). While Taq error may lead to the amplification of artifact copies of ITS, the Taq error rate of *ExTaq* is so low and the frequency of unique clones is so high that Taq error could not account for the level of intragenomic polymorphism encountered in this study. Extensive incomplete concerted evolution in *C. fremontii* is undeniable. Compounding support for this phenomenon are the results of the "more thorough" cloning effort in which 96 clones were sequenced from four individuals of *C. fremontii* (24 clones/individual). It appears the effort was not as exhaustive as anticipated; after sequencing 96 clones from four individuals, an endpoint to the number of unique copies per individual was never reached, nor were issues of paralogy versus orthology illuminated.

The dubious relationships that can arise due to the presence of orthologs and paralogs within the dataset are a major complication when using ITS for phylogenetic study (Alvarez and Wendell 2003). Bailey et al. (2003) suggested gene trees as a way to distinguish benign paralogs (i.e., shallow paralogs) from potentially problematic paralogs (i.e., deep paralogs). Deep paralogs, which are derived from duplications and divergence events prior to speciation, can lead to false phylogenetic inference. The high degree of polyphyly of clones from individuals and from species reflects deep paralogy within North American subg. *Viorna*. Such paralogy is often attributed to introgression and hybridization; however, hybridization is not well documented in wild Clematis. Keener (1967, 1975) suggested *C. coactilis* was a possible result of ancient hybridization between *C. ochroleuca* and *C. albicoma*. Our data suggest an opposite scenario in that if hybridization did occur, it was likely between *C. ochroleuca* and *C. coactilis*. Individuals of *C. ochroleuca* comprise a monotypic clade and individuals of *C. coactilis* comprise two monotypic clades, whereas *C. albicoma* displays a low level of species distinction with clones from individuals scattered throughout the tree and seldom grouped. Though it cannot be ruled out, hybridization is an unlikely explanation for the irresolution between all species because with the exception of *C. ochroleuca*, subsect. *Integrifoliae* taxa occur in isolated populations with little chance of gene flow.

While the topology of the phylogenetic tree suggests deep paralogy and incomplete concerted evolution, incomplete lineage sorting within North American subg. *Viorna* cannot be ruled out. Since all of the observed alleles coalesce to a point within Clade V, it appears that incomplete lineage sorting spans at least as far back as the most common ancestor to North American subg. *Viorna*. However, this conclusion must be made with caution as only a small

portion of this large and diverse genus was included in the gene tree, with the exception of subsect. *Integrifoliae*.

Alternately, non-monophyly of species due to a lack of informative genetic variation can be indicative of recent divergence. The notion that these taxa could be recently diverged is supported by the conclusions of Miikeda et al. (2006) who suggested a recent rapid radiation of Clematidinae subsequent to divergence from Anemoniae. Recent divergence explains the lack of resolution found within the tree, as this implies inadequate time to reach completion of concertion.

Although the majority of the tree (Figure 4) is unresolved, two small clades (S and E) have more structure than the others (Figure 5). Interestingly, these clades both contain pseudogenes. Within clade S is a tertiary clade comprised solely of putative pseudogenes from five different individuals of *C. fremontii*. Sister to the clade of identified pseudogenes is a clade containing one individual (FMO1.602) from population FMO1 and one individual (FMO5.672) from FMO5 and sister to all of these is a clade containing clones from *C. albicoma* and *C. viticaulis*. The well-defined pseudogene clade embedded within clade S is interesting as it is much more resolved than any of the other clades (as would be expected for a gene released from functional constraint). What is also interesting is that this level of resolution was maintained back to the beginning of Clade S, suggesting that Clade S represents a pseudogene that was derived prior to the separation of species (since clones from *C. albicoma* and *C. viticaulis* resolve to this clade). An alternative explanation is that the pseudogene actually formed at the point marked with the large red dot and that the putative functional copies within Clade S (some from *C. fremontii, C. albicoma, and C. viticaulis*) represent precursors to pseudogene formation. The increased level of resolution in this clade allows for phylogeographic inferences about some of the *C. fremontii* populations. A clone from an individual (FTN1.570) in Tennessee is sister to a clone from an individual (FMO5.669) from the southernmost Missouri population (FMO5) and to an individual (FMO3.660) from a more northern Missouri population (FMO3).

The only other clade with an equivalent level of resolution to Clade S is Clade E (Figure 5). A pseudogene belonging to a different individual (FMO5.624) from population FMO5 is located in a quaternary clade within Clade E; this coupled with the increased resolution of this clade suggests that it too is a clade of pseudogenes. Curiously, relationships shown in this clade support some of those in Clade S in that a clone from the southernmost Missouri population is

sister to a clone from the same individual from Tennessee (FTN.570) that is found in Clade S. These placements indicate the Tennessee population and the southern Missouri population are closely related. This relationship is supported by the topography and associate species of the southern Missouri population, which more closely resemble the Tennessee population than the other more northern populations in Missouri (i.e., the habitats are similar in appearance).

Although, Clade E does not include a well resolved pseudogene grouping like clade S, the high level of structure in both implies that nrDNA copies within the individuals in these clades are able to accumulate more mutations than the majority of the other copies sampled in this study. The inclusion of a putative pseudogene within Clade E also suggests that it too could include pre-pseudogenic copies. The inclusion of pseudogenes in phylogenetic analysis is controversial (e.g. Bailey et al. 2003, Harpke and Peterson 2006, Alvarez and Wendel 2003, Razafimandimbison et al. 2004). However this study supports the notion that in instances of low resolution, pseudogenes can provide the necessary level of nucleotide diversity for phylogenetic inference *at the population level*, since inference at higher levels could lead to false relationships.

Ultimately, the intraindividual polymorphism is so vast within North American subg. *Viorna* and the level of polyphyly among subgenera is so high that the use of ITS at any level of study in the Clematidinae below the level of genus is impractical. The cloning effort that would be necessary to parse out lineage relationships is inconceivable, as demonstrated by the "exhaustive" cloning effort for which no endpoint to the number of unique copies within an individual was reached. The phenomena of incomplete concerted evolution and widespread polyphyly haunt lineages of subg. *Viorna* to such an extent that future use of ITS in *Clematis* for phylogenetic inference is discouraged.

LIST OF REFERENCES

Ainouche, M.L., and R.J. Bayer. 1997. On the origins of the tetraploid *Bromus* species (section *Bromus*, Poaceae): insights from internal transcribed spacer sequences of nuclear ribosomal DNA. Genome 40(5): 730-743.

Alvarez, I. and J.F. Wendel. 2003. Ribosomal ITS sequences and plant phylogenetic inference. Molecular Phylogenetics and Evolution 29: 417-434.

APG II. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. Botanical Journal of the Linnean Society 141: 339-436.

Bailey, C.D., T.G. Carr, S.A. Harris, and C.E. Hughes. 2003. Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes. Molecular Phylogenetics and Evolution 29: 435-455.

Baldwin, B.G., M.J. Sanderson, J.M. Porter, M.F. Wojciechowski, C.S. Campbell, and M.J. Donoghue, 1995. The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. Ann. Mo. Bot. Gard. 82, 247–277.

Bennett, M.D. and J.B. Smith. 1991. Nuclear DNA Amounts in Angiosperms. Philosophical Transactions: Biological Sciences. 334: 309-345.

Buckler E.S. IV, A. Ippolito and T.P. Holtsford. 1997. The Evolution of Ribosomal DNA: Divergent Paralogues and Phylogenetic Implications. Genetics. 145: 821-832.

Devos, N., S. Oh, O. Raspé, A. Jacquemart, and P.S. Manos. 2005. Nuclear ribosomal DNA sequence variation and evolution of spotted marsh-orchids (*Dactylorhiza maculate* group). Molecular Pylogenetics and Evolution 36: 568-580.

Douzery, E.J.P., A.M. Pridgeon, P. Kores, H.P.Linder, H. Kurzweil, M.W.Chase. 1999. Molecular phylogenetics of *Diseae* (Orchidaceae): a contribution from nuclear ribosomal ITS sequences. Am. J. Bot. 86: 887-899 .

Dover, G. 1982. Molecular drive: a cohesive mode of species evolution. Nature. 299(5879): 111-117.

Elder, J.F. Jr. and B.J. Turner. 1995. Concerted Evolution of Repetitive DNA Sequences in Eukaryotes. The Quarterly Review of Biology. 70(3): 297-320.

Erickson, R.O. 1943. Taxonomy of *Clematis* section *Viorna*. Annals of the Missouri Botanical Garden 30(1): 1-62.

ExTaq Polymerase. TaKaRa Bio USA. www.takara-bio.us/docs/TAK_RR001_FAQ.pdf.

Feliner, G.N, and J.A. Rosselló. 2007. Better the devil you know? Guidelines for insightful utilization of nrDNA ITS species-level evolutionary studies in plants. Molecular Phylogenetics and Evolution 44: 911-919.

Feliner, G.N, B.G. Larenza, and J.F. Aguilar. 2004. Fine-scale Geographical Structure, Intra-individual Polymorphism and Recombination in Nuclear Ribosomal Internal Transcribed Spacers in *Armeria* (Plumbaginaceae). Annals of Botany 93: 189-200.

Gregory, W.C. 1941. Phylogenetic and Cytological Studies in the Ranunculaceae Juss. Transactions of the American Philosophical Society. 31(5) 443-521.

Grey-Wilson. 2000. *Clematis*, the genus: a comprehensive guide for gardeners, horticulturists, and botanists.

Harpke, D. and A. Peterson. 2006. Non-concerted ITS evolution in *Mammillaria* (Cactaceae). Molecular Phylogenetics and Evolution. 41(3) 579-593.

Harpke, D. and A. Peterson. 2007. Quantitative PCR Revealed a Minority of ITS Copies to Be Functional in *Mammillaria* (Cactaceae). Int. J Plant Sci. 168(8): 1157–1160.

Harpke, D. and A. Peterson. 2008. Note: 5.8S motifs for the identification of pseudogenic ITS regions. Botany 86: 300-305.

Hillis, D.M. and M.T. Dixon. 1991. Ribosomal DNA: Molecular Evolution and Phylogenetic Inference. The Quarterly Review of Biology. 66(4): 411-453.

Hughes, C. E., C. D. Bailey, S. Krosnick, and M. Luckow. 2003. Relationships among genera of the informal *Dichrostachys* and *Leucaena* groups (Mimosoideae) inferred from nuclear ribosomal ITS sequences. In B. Klitgaard and A. Bruneau [eds.], Advances in Legume Systematics 10, Pgs 221-238, Royal Botanic Gardens, Kew.

Keener, C. S. 1967. A biosystematic study of *Clematis* subsection *Integrifoliae* (Ranunculaceae). J. Elisha Mitchell Sci. Soc. 83: 1-40.

Keener, C. S. 1975. Studies in the Ranunculaceae of the Southeastern United States. I. *Anemone* L. Castanea 40(1) 36-44.

Kim K.J. and R.K. Jansen. 1994. Comparisons of phylogenetic hypotheses among different data sets in dwarf dandelions (*Krigia*, Asteraceae): additional information from internal transcribed spacer sequences of nuclear ribosomal DNA. Pl. Syst. Evol. 190(3-4): 157-185.

King, M.G. and E.H. Roalson. 2008. Exploring Evolutionary Dynamics of nrDNA in *Carex* Subgenus *Vignea* (Cyperaceae). Systematic Botany 33(3): 514-524.

Kita, Y. and M. Ito. 2000. Nuclear ribosomal ITS sequences and phylogeny in East Asian *Aconitum* subgenus *Aconitum* (Ranunculaceae), with special reference to extensive polymorphism in individual plants. Plant Systematics and Evolution. 255: 1-13.

Koch, M.A., C. Dobe, T. Mitchell-Olds. 2003. Multiple Hybrid Formation in Natural Populations: Concerted Evolution of the Internal Transcribed Spacer of Nuclear Ribosomal DNA (ITS) in North American *Arabis divaricarpa* (Brassicaceae). Mol. Biol. Evol. 20(3): 338-350.

Long, E.O. and I.B. Dawid. 1980. Repeated genes in eukaryotes. Ann. Rev. Biochem. 49: 727-764.

Miikeda, O., K. Kita, T. Handa, T. Yukawa. 2006. Phylogenetic relationships of *Clematis* (Ranunculaceae) based on chloroplast and nuclear DNA sequences. Botanical Journal of the Linnean Society 152: 153–168.

Morgan, D. R. 2003. nrDNA external transcribed spacer (ETS) sequence data, reticulate evolution, and the systematics of *Machaeranthera* (Asteraceae). Systematic Botany 28: 179-190.

Noyes, R.D. 2006. Intraspecific nuclear ribosomal DNA divergence and reticulation in sexual diploid *Erigeron strigosus* (Asteraceae). American Journal of Botany 93(3): 470-476.

Okuyama, Y., N. Fujii, M. Wakabayashi, A. Kawakita, M. Ito, M. Watanabe, N. Murakami, and M. Kato. 2005. Nonuniform concerted evolution and chloroplast capture: heterogeneity of observed introgression patterns in three molecular data partition phylogenies of Asian *Mitella* (Saxifragaceae). Molecular Biology and Evolution 22: 285–296.

Posada, D. 2005. Collapse ver. 1.2: Describing haplotypes from sequence alignments. Available at URL: http://www.darwin.uvigo.es/software/collapse.html.

Quijada A., A. Liston, W. Robinson, and E. Alvarez Buylla. 1997. The ribosomal ITS region as a marker to detect hybridization in pines. Molec. Ecol. 6. (10): 995-996.

Rauscher J.T., J.J. Doyle, and H.D. Brown. 2004. Multiple Origins and nrDNA Internal Transcribed Spacer Homeologue Evolution in the *Glycine tomentella* (Leguminosae) Allopolyploid Complex. Genetics. 166: 987-998.

Razafimandimbison, S.G., E.A. Kellogg, and B. Bremer. 2004. Recent origin and phylogenetic utility of divergent ITS putative pseudogenes: A case study from *Naucleeae* (Rubiaceae). Syst. Biol. 53(2):177-192.

Rieseberg, L.H., H.C. Choi, and D. Ham. 1991. Differential Cytoplasmic versus Nuclear Introgression in *Helianthus*. The Journal of Heredity. 82(6): 489-493.

Saini, A., S.K. Reddy, N. Jawali. 2008. Intra-individual and intra-species heterogeneity in nuclear rDNA ITS region of *Vigna* species from subgenus *Ceratotropis*. Genetics Research. 90: 299-316.

Sang, T, D.J. Crawford, and T.F. Stuessy. 1995. Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: Implications for biogeography and concerted evolution. Proc. Natl. Acad. Sci. 92: 6813-6817.

Small, R. L., R. C. Cronn, and J. F. Wendel. 2004. L. A. S. JOHNSON REVIEW No. 2. Use of nuclear genes for phylogeny reconstruction in plants. Australian Systematic Botany 17: 145–170.

Swofford, D. L. 2002. PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4.0b10. Sinauer, Sunderland, Massachusetts, USA.

Takezaki, N., A. Razhetsky, and M. Nei. 1995. Phylogenetic test of the molecular clock and linearized trees. Mol. Biol. Evol. 12: 823-833.

Wagner, A., N. Blackstone, P. Cartwright, M. Dick, B. Misof, P. Snow, G.P. Wagner, J. Bartels, M. Murtha, and J. Pendleton. 1994. Surveys of gene families using polymerase chain reaction—PCR selection and PCR drift. Systematic Biology. 43: 250-261.

Weeks, A. and B.B. Simpson. 2004. Molecular genetic evidence for interspecific hybridization among endemic Hispaniolan *Bursera* (Burseraceae). American Journal of Botany. 91: 976-984.

Wei, X.X., X.Q. Wang, and D.Y. Hong. 2003. Marked Intragenomic Heterogeneity and Geographical Differentiation of nrDNA ITS in *Larix potaninii* (Pinaceae). Journal of Molecular Evolution. 57(6): 623-625.

Wendel, A., A. Schabel, and T. Seelanan. 1995. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (Gossypium). PNAS. 92(1): 280-284.

White, T.J., Bruns, T., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M., Gelfand, D., Sninsky, J., White, T. (Eds.), PCR Protocols: A Guide to Methods and Applications. Academic Press,San Diego, pp. 315–322.

Wichman, S.R., Wright, S.D., Cameron, E.E., Keeling, D.J., and R.C. Gardner. 2002. Elevated genetic heterogeneity and Pleistocene climatic instability: inferences from nrDNA in New Zealand *Coprosma* (Rubiaceae). Journal of Biogeography. 29(7): 943-954.

Yang, T.Y.A. and D.M. Moore. 1999. A Revision of the Viorna Group of Species (Section Viorna sensu Prantl) in the Genus *Clematis* (Ranunculaceae). Systematics and Geography of Plants. 68(1/2): 281-303.

Table 1. Taxa used in this investigation, source and voucher numbers, population name, total DNA number, and GenBank accession numbers. First letter of population name represents the species; second and third letters represent the state in which population is located, number distinguishes multiple populations per state according to order of population sampling.

Species	Collector; Number; Year	Herb.	State and County	Population	DNA #	ITS1-5.8S-ITS2 GenBank #
C. fremontii S. Watson	C. Montgomery; CMo1; 2007	UCHT	USA; TN; Hamilton	FTN1		
					563	
					564	*
					565	*
					566	*
					567	*
					568	*
					569	*
					570	*
					571	*
					572	*
					573	*
					574	*
					575	*
					576	
					577	*
					578	*
					579	*
					580	*
					581	*
					582	*
					583	
C. fremontii S. Watson			USA; GA; Floyd	FGA1		
					584	
					585	*
					586	*
					587	*
					588	*
					589	*
					590	*
					591	*
					592	*
					593	*
					594	*
					595	*
					596	*
					597	*
C. fremontii S. Watson	C. Montgomery; CMo2; 2008	UCHT	USA; MO; Jefferson	FMO1		
					601	*
					602	*
					603	*
					604	*
					605	*

Table 1. Taxa used in this investigation, source and voucher numbers, population name, total DNA number, and GenBank accession numbers. First letter of population name represents the species; second and third letters represent the state in which population is located, number distinguishes multiple populations per state according to order of population sampling.

Species	Collector; Number; Year	Herb.	State and County	Population	DNA #	ITS1-5.8S-ITS2 GenBank #
C. fremontii S. Watson	C. Montgomery; CMo3; 2008	UCHT	USA; MO; Franklin	FMO2		
					606	*
					607	*
					608	*
					609	*
					610	*
					653	*
					654	*
					655	*
					656	*
					657	*
C. fremontii S. Watson	C. Montgomery; CMo4; 2008	UCHT	USA; MO; Jefferson	FMO3		
					611	*
					612	*
					613	*
					614	*
					615	*
					658	*
					659	*
					660	*
					661	*
					662	*
C. fremontii S. Watson	C. Montgomery; CMo5; 2008	UCHT	USA; MO; Washington	FMO4		
					616	*
					617	*
					618	*
					619	*
					620	*
					663	*
					664	*
					665	*
					666	*
					667	*
C. fremontii S. Watson	C. Montgomery; CMo6; 2008	UCHT	USA; MO; Ozark	FMO5		
					621	*
					622	*
					623	*
					624	*
					625	*
					668	*
					669	*
					670	*
					671	*
					672	*

Table 1. Taxa used in this investigation, source and voucher numbers, population name, total DNA number, and GenBank accession numbers. First letter of population name represents the species; second and third letters represent the state in which population is located, number distinguishes multiple populations per state according to order of population sampling.

Species	Collector; Number; Year	Herb.	State and County	Population	DNA #	ITS1-5.8S-ITS2 GenBank #
C. fremontii S. Watson	C. Montgomery; CMo7; 2008	UCHT	USA; KS; Cloud	FKS1		
					626	*
					627	*
					628	*
					629	*
					630	*
					673	*
					674	*
					675	
					676	*
					677	*
C. fremontii S. Watson	C. Montgomery; CMo8; 2008	UCHT	USA; KS; Mitchell	FKS2		
					631	*
					632	*
					633	*
					634	*
					635	*
					678	*
					679	*
					680	*
					681	*
					682	*
C. fremontii S. Watson	C. Montgomery; CMo9; 2008	UCHT	USA; KS; Rooks	FKS3		
					636	*
					637	*
					638	*
					639	*
					640	*
					683	*
					684	*
					685	*
					686	*
					687	*
C. fremontii S. Watson	C. Montgomery; CMo10; 2008	UCHT	USA; KS; Phillips	FKS4		
					641	*
					642	*
					643	*
					644	*
					645	*
					688	*
					689	*
					690	*
					691	*
					692	*

Table 1. Taxa used in this investigation, source and voucher numbers, population name, total DNA number, and GenBank accession numbers. First letter of population name represents the species; second and third letters represent the state in which population is located, number distinguishes multiple populations per state according to order of population sampling.

Species	Collector; Number; Year	Herb.	State and County	Population	DNA #	ITS1-5.8S-ITS2 GenBank #
C. fremontii S. Watson	C. Montgomery; CMo11; 2008	UCHT	USA; KS; Jewell	FKS5		
					646	*
					647	*
					648	*
					649	*
					650	*
					693	*
					694	*
					695	*
					696	*
					697	*
C. viorna L.	C. Montgomery; CMo12; 2008	UCHT	USA; NC; Jackson	RNC1		
					600	*
C. glaucophylla Small	G. Woodworth; GWo-18; 2008	UCHT	USA; AL; Bibb	GAL1		
					699	
C. addisonii Britton	C. Montgomery; CMo16; 2009	UCHT	USA; VA; Montgomery	DVA1		
					700	*
C. coactilis (Fernald) Keener	C. Montgomery; CMo13; 2009	UCHT	USA; VA; Craig	CVA1		
					701	*
					702	*
					703	*
					704	*
					705	*
C. viticaulis Steele	C. Montgomery; CMo14; 2009	UCHT	USA; VA; Bath	VVA1		
					706	*
					707	*
					708	*
					709	*
					710	*
C. albicoma Wherry	C. Montgomery; CMo15; 2009	UCHT	USA; VA; Alleghany	AVA1		
					711	*
					712	*
					713	*
					714	*
					715	*
C. ochroleuca Aiton	C. Montgomery; CMo17; 2009	UCHT	USA; GA; Franklin	OGA1		
					721	*
					722	*
					723	
					724	*
					725	

Table 2. Primers used for PCR amplification and sequencing

nrDNA Region	Primer name and sequence (5'-3')
ITS1-5.8S-ITS2	**AB101:** ACG AAT TCA AGG TCC GGT GAA GTG TTC G (Douzery et al., 1999)
	AB102: TAG AAT TCC CCG GTT CGC TCG CCG TTA C (Douzery et al., 1999)

Table 3. List of haplotypes of ITS clones. Clones are labeled by capital letters and listed by accession. Assigned haplotype number is listed per clone. Species and populations are listed above corresponding accessions. Populations containing accessions are represented by population labels from Table 1. Gray cells indicate shared haplotypes assigned to more than one clone. Putative [seudogenes are indicated with a slash through the haplotype number. The DNA numbers of individuals chosen for thorough cloning are in bold.

			No. of haplotypes per individual	Haplotype number of ITS clones																								
Species	Population	DNA #		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
C. fremontii S. Watson	FTN1																											
		563	4	186	187	188	189																					
		568	4	190	191	192	193																					
		570	18	194	195	196	16	287	200	16	288	200	200	200	~~289~~	290	291	292	24	200	293	294	2	16	295	~~296~~	297
		572	4	197	~~198~~	199	200																					
		575	3	200	201	200	5																					
		577	4	5	202	203	204																					
		582	4	205	206	24	207																					
C. fremontii S. Watson	FGA1																											
		587	4	1	2	3	4	4																				
		588	4	5	6	7	8																					
		590	18	9	2	10	2	298	299	300	24	2	24	301	302	4	303	171	4	24	304	47	305	24	306	307	308
		594	3	11	12	11	13																					
		595	4	14	15	16	17																					
C. fremontii S. Watson	FMO1																											
		601	7	16	98	99	100	101	102	103																		
		602	4	104	105	106	107																					
		603	4	2	108	24	109																					
		604	15	47	110	47	111	309	47	310	47	311	312	47	313	5	314	5	47	24	315	313	316	47	317	47	318
		605	4	112	113	114	115																					
C. fremontii S. Watson	FMO2																											
		607	6	116	117	334	118	119	120																			
		608	4	121	122	5	123																					
		610	4	124	47	24	125																					
		653	4	~~126~~	127	128	129																					
		655	3	2	5	130	2	5	130																			
C. fremontii S. Watson	FMO3																											
		611	3	4	2	4	4	131																				
		612	5	24	132	133	134	135																				
		614	5	2	2	136	137	2	138	2	139																	
		615	7	140	141	142	2	143	144	145																		
		660	5	146	~~147~~	47	146	148	~~147~~	47	149																	
C. fremontii S. Watson	FMO4																											
		617	4	150	24	151	152																					
		618	4	153	154	155	156																					
		619	4	47	~~157~~	2	158																					
		665	4	159	160	24	161																					
		666	4	~~162~~	163	164	5																					
C. fremontii S. Watson	FMO5																											
		624	5	165	166	167	168	169																				
		669	3	170	171	171	~~172~~																					
		670	4	173	24	174	2																					
		671	7	175	176	177	24	24	178	179	2																	
		672	8	4	180	181	~~182~~	183	184	5	185																	
C. fremontii S. Watson	FKS1																											
		626	4	18	19	20	2																					
		628	5	21	22	2	23	2	24																			
		673	4	25	26	~~27~~	28																					
		674	4	29	30	31	32																					
		677	4	4	33	34	35	4																				
C. fremontii S. Watson	FKS2																											
		633	21	36	37	38	39	47	~~319~~	320	321	322	2	323	324	325	326	327	328	2	329	330	331	332	47	2	333
		634	4	40	41	2	42																					
		635	4	43	44	45	46																					
		679	4	2	47	5	48																					
		680	3	~~49~~	2	2	50																					
C. fremontii S. Watson	FKS3																											
		636	5	51	52	24	53	54																				
		637	4	55	56	57	~~58~~																					
		638	4	59	24	60	61																					
		639	5	62	63	2	64	65	2																			
		640	5	66	2	67	68	69																				
C. fremontii S. Watson	FKS4																											
		641	3	2	70	71	2																					
		645	4	72	2	73	4																					
		688	4	74	75	76	77																					
		689	3	24	4	24	78																					
		690	4	79	2	80	81																					
C. fremontii S. Watson	FKS5																											
		646	2	2	2	82	2																					
		649	5	83	84	85	86	87																				
		650	6	5	88	89	90	91	92																			
		695	4	93	16	5	24																					
		697	4	94	95	96	97																					

Table 3. List of haplotypes of ITS clones. Clones are labeled by capital letters and listed by accession. Assigned haplotype number is listed per clone. Species and populations are listed above corresponding accessions. Populations containing accessions are represented by population labels from Table 1. Gray cells indicate shared haplotypes assigned to more than one clone. Putative [seudogenes are indicated with a slash through the haplotype number. The DNA numbers of individuals chosen for thorough cloning are in bold.

			No. of haplotypes per individual	Haplotype number of ITS clones																								
Species	Population	DNA #		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
C. coactilis (Fernald) Keener	CVA1																											
		701	3	228	229	47	229																					
		702	4	230	231	232	233																					
		703	5	~~234~~	~~235~~	236	~~237~~	47																				
		704	6	238	239	240	241	242	242	243																		
		705	5	244	245	246	47																					
C. viticaulis Steele	VVA1																											
		706	4	272	273	274	275																					
		707	4	276	277	278	279																					
		708	3	280	281	24	24																					
		709	3	211	211	282	283																					
		710	4	284	285	286	47																					
C. albicoma Wherry	AVA1																											
		711	4	211	211	24	212	213																				
		712	4	211	214	215	216																					
		713	4	217	218	219	220																					
		714	5	221	211	222	~~223~~	224																				
		715	6	225	155	226	24	24	227	5																		
C. ochroleuca Aiton	OGA1																											
		721	4	251	252	253	254	254	251																			
		722	4	251	255	256	257																					
		723	4	258	259	260	47																					
		724	6	261	262	210	263	264	265																			
		725	4	266	266	267	266	268	269																			
C. viorna L.	RNC1																											
		600	4	24	270	47	271																					
C. glaucophylla Small	GAL1																											
		699	3	208	209	210	cpDNA																					
C. addisonii Britton	DVA1																											
		700	4	247	248	249	250																					

Table 4. List of *Clematis* taxa downloaded from GenBank, including subgeneric rank of taxa, if available, and corresponding accession numbers.

Species	Subgenus	GenBank accession number
C. alternata Kitamura & Tamura	*Campanella* Tamura	AB120190
C. japonica Thunb.	*Campanella* Tamura	AB120187
C. lasiandra Maxim.	*Campanella* Tamura	AB120185
C. nobilis Nakai	*Campanella* Tamura	AB120206
C. ochotensis (Pall.) Poir.	*Campanella* Tamura	AB120182
C. orientalis L.	*Campanella* Tamura	AB120196
C. serratifolia Rehd.	*Campanella* Tamura	AB120205
C. stans Sieb. & Zucc.	*Campanella* Tamura	AB120188
C. tangutica (Maxim.) Korsh.	*Campanella* Tamura	AB120195
C. villosa DC.	*Campanella* Tamura	AB120211
C. afoliata J.Buch.	*Clematis* Spach	AB120193
C. apiifolia DC.	*Clematis* Spach	AB120180
C. eichleri Tamura	*Clematis* Spach	AB120209
C. fasciculiflora Franch.	*Clematis* Spach	AB120203
C. gentianoides DC.	*Clematis* Spach	AB120210
C. lasiantha Nutt. ex Torr. & A.Gray	*Clematis* Spach	AB120200
C. ligusticifolia Nutt. ex Torr. & A.Gray	*Clematis* Spach	AB117593
C. pierotii Miq.	*Clematis* Spach	AB120191
C. potaninii Maxim.	*Clematis* Spach	AB120198
C. tashiroi Maxim.	*Clematis* Spach	AB120192
C. vitalba L.	*Clematis* Spach	AB120207
C. williamsii A.Gray	*Clematis* Spach	AB120181
C. angustifolia Jacq.	*Flammula DC.*	AB120199
C. brachyura Maxim.	*Flammula DC.*	AB120204
C. crassifolia Benth.	*Flammula DC.*	AB120194
C. delavayi Franch.	*Flammula DC.*	AB120202
C. florida Thunb.	*Flammula DC.*	AB120186
C. hexapetala Pall.	*Flammula DC.*	AY055385
C. patens Morr. & Decne.	*Flammula DC.*	AB120184
C. terniflora DC.	*Flammula DC.*	AB120183
C. uncinata Champ. var. *ovatifolia* (T.Ito) Ohwi	*Flammula DC.*	AB120189
C. fusca Turcz.	*Viorna* (Rchb.) Tamura	AB120179
C. texensis Buckley	*Viorna* (Rchb.) Tamura	AB120197
C. armandii Franch.		FJ572047
C. gratopsis W. T. Wang		FJ424226
C. javana DC.		DQ499136
C. montana Buchanan-Hamilton ex DC.		FJ424227
C. paniculata J. F. Gmel.		DQ499135

Table 5. List of shared haplotypes. Assigned haplotype number, total number of clones per haplotype and total number of accessions per haplotype are listed. Number of clones displaying haplotype are listed by population. In instances of haplotype sharing by multiple clones from one accession, the number of individuals displaying the haplotype is listed in parentheses.

			Population																		
Haplotype #	Clones per haplotype	Accessions per haplotype	FTN1	FGA1	FMO1	FMO2	FMO3	FMO4	FMO5	FKS1	FKS2	FKS3	FKS4	FKS5	CVA1	VVA1	AVA1	OGA1	RNC1	GAL1	DVA1
11	2	1		2																	
130	2	1				2															
146	2	1					2														
147	2	1					2														
229	2	1													2						
242	2	1													2						
254	2	1																2			
313	2	1			2																
266	3	1																3			
155	2	2						1									1				
210	2	2																1		1	
171	3	2		1					2(1)												
251	3	2																3(2)			
200	8	3	8(3)																		
16	6	4	3(1)	1	1									1							
211	6	4														2(1)	4(3)				
4	11	7		4(2)			3(1)		1	2(1)			2								
5	15	11	2		2(1)	3(2)		1	1		1			2			1				
47	22	13		1	8(1)	1	2(1)	1			3(2)				3	1		1	1		
24	27	20	2	4(1)	2	1	1	2	3(2)	1		2	2(1)	1		2(1)	3(2)		1		
2	37	23	1	4(2)	1	2(1)	6(3)	1	2	3(2)	7(4)	3(2)	4(3)	3(1)							

Figure 1. Nuclear ribosomal genes, featuring the ITS1-5.8S-ITS2 region. Nuclear ribosomal genes are arranged in hundreds to thousands of tandem repeats in one or more arrays (upper depiction). Each repeat (lower depiction) contains three exons (18S, 5.8S and 26S) and two introns (ITS1 and ITS2). The exons code for the rRNA molecules found in the large and small ribosomal subunits and are separated by ITS1 and ITS2. The internal transcribed spacer sequences (ITS1 and ITS2) are cleaved from precursor transcripts during formation of the mature rRNA. Primer binding sites (ITSAB101 and ITSAB102) are indicated by directional arrows.

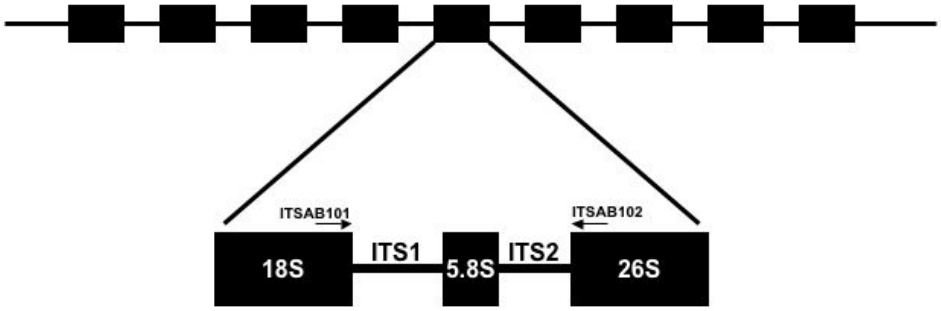


Figure 2. Map of C. fremontii and subsection Integrifoliae populations and subsect. Viorna outgroup taxa sampled in this study.

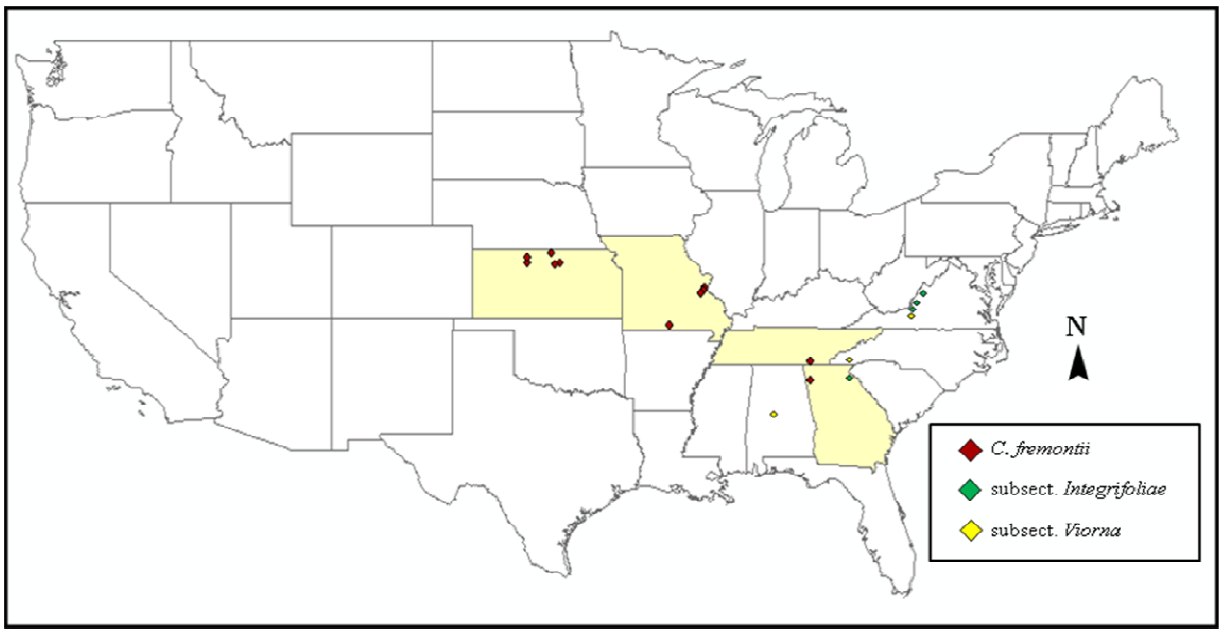
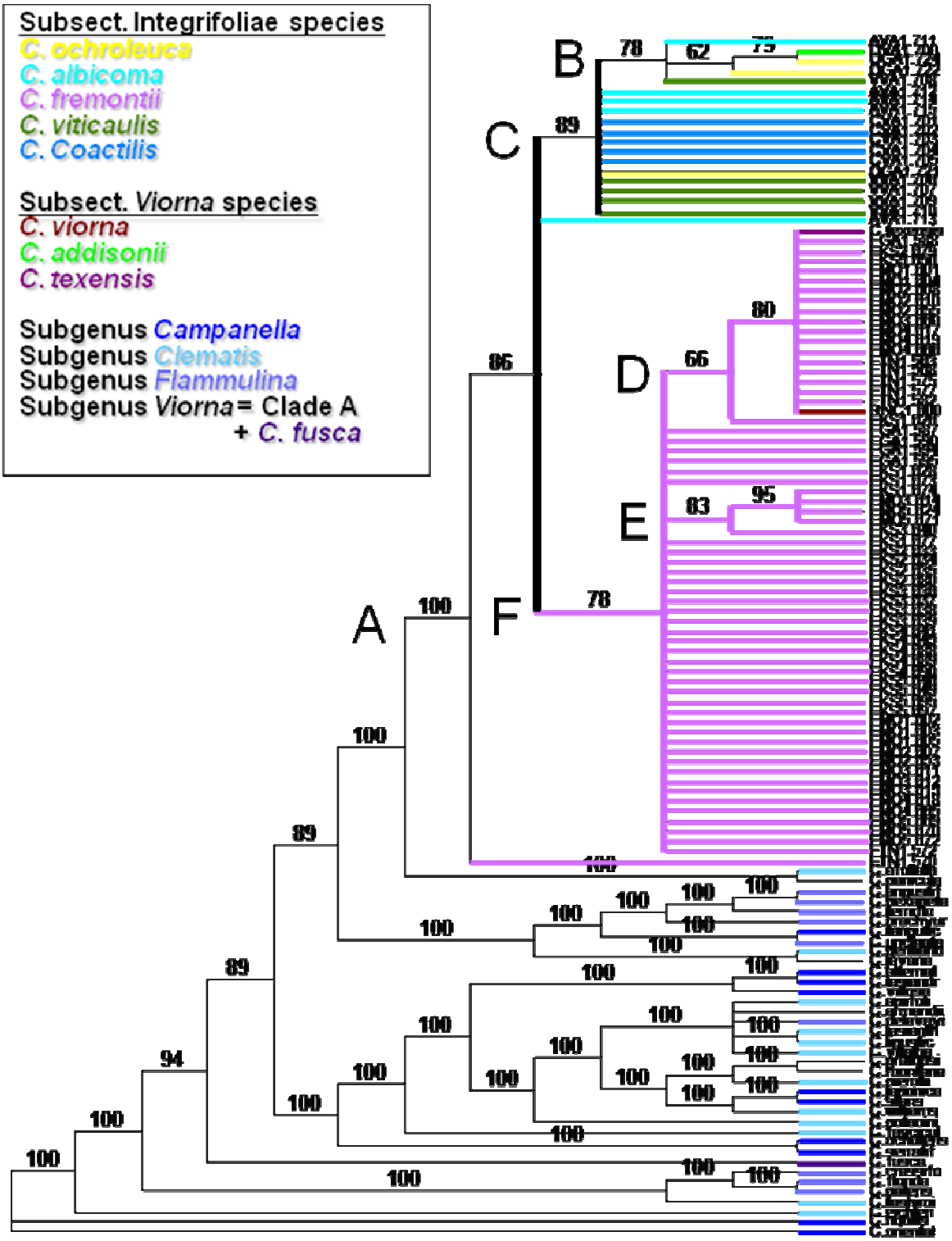


Figure 3. 50% Majority rule consensus tree of ITS sequences from direct PCR dataset.



.

Figure 4. 50% Majority rule consensus tree of ITS sequences from clone dataset.

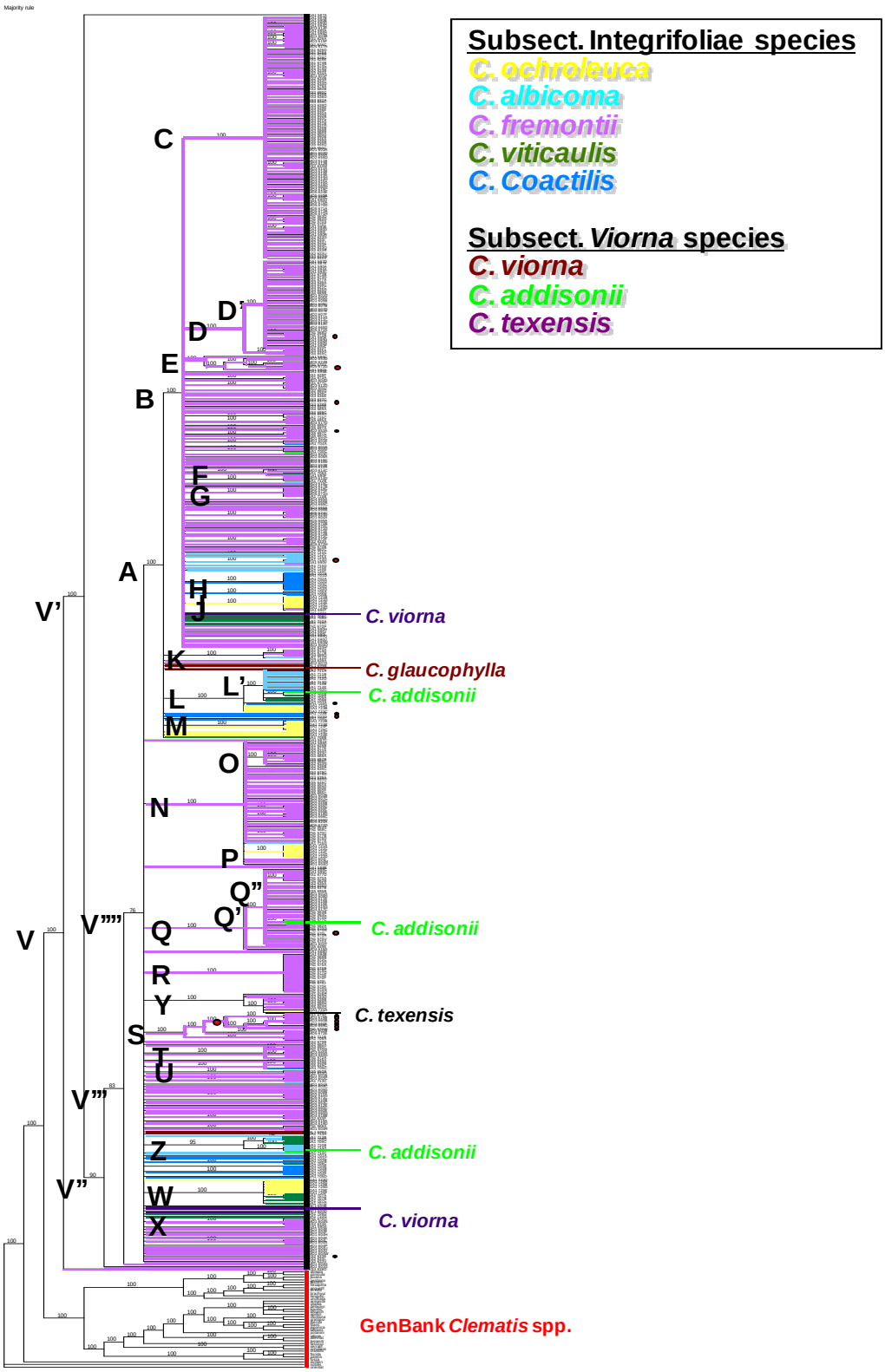
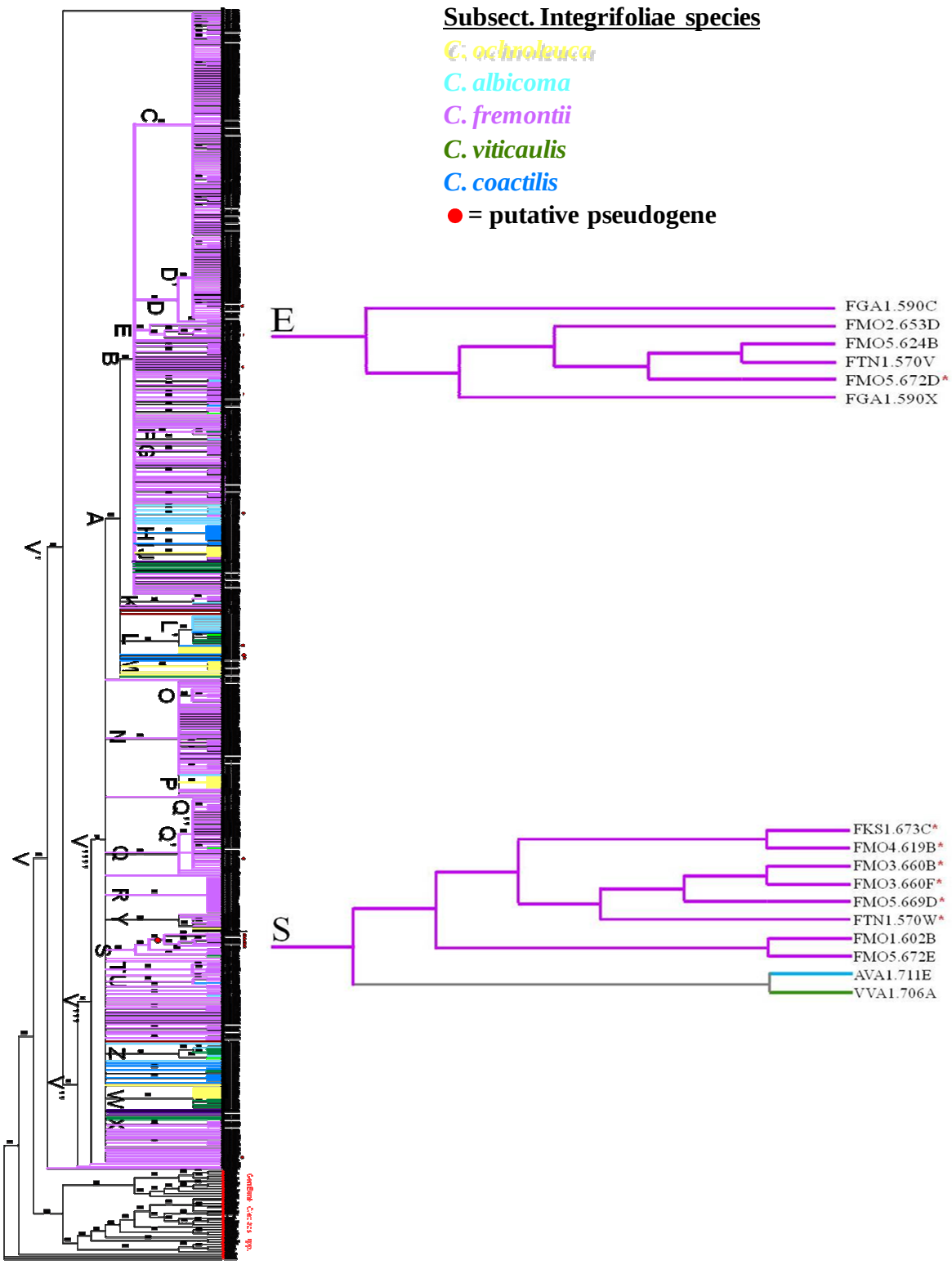


Figure 5. Clades E and S from 50% Majority rule consensus tree (Figure 4) of ITS sequences from clone dataset



Part 4. The Single Copy Nuclear Gene *G3pdh*: A Promising Marker for a Phylogeographic Investigation of *C. fremontii*

INTRODUCTION

Traditional approaches in plant population genetics utilize selectively neutral, unordered molecular markers that are highly variable, such as allozymes and microsatellites. These are attractive methods for population genetics because the observed variation distributed among populations is believed to represent gene flow and genetic drift (Schaal et al. 1998). Though allozyme and microsatellite research is beneficial for measuring intrapopulation variation and assessing the distribution of intraspecific variation among populations, these approaches have limitations with respect to placing populations of a species into a phylogenetic (historically ordered) framework (Schaal et al. 1998). In other words, the implementation of these traditional unordered markers in population studies leaves the genealogical relationships among allelic variants uncharacterized. A phylogeographic approach to studying the geographical distribution of genealogical lineages (Avise 2000) offers an alternative method of assessing plant populations in that a gene (or several genes) is sequenced for multiple accessions in several populations to reveal their evolutionary relatedness. Through analyzing genealogically ordered sequence data, historical perspective of population-level relationships may be revealed.

Phylogeographic studies largely rely on chloroplast and nuclear ribosomal DNA sequences to investigate population level relationships. Although cpDNA and nrDNA are commonly used genomes in plant phylogenetics, both have characteristics that render their use for phylogeographic studies less than desirable. The chloroplast genome is uniparentally inherited (typically maternally in angiosperms); as such cpDNA only traces the lineage of one parent, and thus lacks the genealogical patterns necessary to capture the entire population history (Sang 2002, Hare 2001). Nuclear ribosomal DNA presents a contrasting problem related to the fact that it is a high copy number gene and perhaps too often assumed to be homogenized to one (or few) allele(s) through concerted evolution. Despite nrDNA being biparentally inherited, sequences can become directionally homogenized through concerted evolution, ultimately masking an individual's parental lineage (Alvarez and Wendel 2003). Moreover, the complex architecture and evolutionary mechanisms of nrDNA can (1) result in pseudogene formation, (2) render this gene region difficult to interpret because of paralogy and incomplete lineage sorting, and (3) ultimately make the separation of parental alleles an arduous or even impossible process (Sang 2002, Small et al. 2004, Montgomery and Shaw unpublished data). While noncoding cpDNA and nrDNA can be informative at the interspecific level, both rarely diverge quickly

enough to elucidate intraspecific or population-level relationships in plants (Small et al. 1998, Balwin et al. 1995).

In instances where historically ordered datasets are necessary and cpDNA and/or nrDNA have yielded insignificant resolution or conflicting results, many researchers have looked to low-copy nuclear genes for increased resolution. Small (et al. 2004, p. 165) went so far as to claim that single copy nDNA appears to be the only reasonable option when attempting to parse out ancestral lineage in instances of "temporally compressed divergence events" such as recent radiation or hybridization.

Single copy nuclear genes are particularly useful molecular tools because they have quickly evolving introns and occur in only two possible allelic forms in diploid species; therefore, they should contain ample molecular variation to infer population relationships, and they should not be complicated by paralogy. These features enhance the ease of capturing and separating all of an individual's ancestral lineages, which is critical in low-level molecular studies.

Comprised of about 280 species, *Clematis* L., has a worldwide distribution and is among the largest genera in Ranunculaceae (Yang and Moore 1999). Ranunculaceae is considered by most botanists to be among the most basal groups of eudicots (APG II, 2003) and is primarily characterized by herbaceous plants with a multi-carpelled gynoecium (Grey-Wilson 2000). Within Ranunculaceae, *Clematis* is characterized by single-seeded achenes, numerous stamens, and an apetaloid corolla (Grey-Wilson 2000). Chromosomal counts of 42 different *Clematis* species revealed that *Clematis* is a diploid genus with eight pairs of chromosomes (2n=16) (Gregory 1941). Polyploidy is rare in *Clematis* but has been documented in three hybrid cultivars (Erickson 1943); however polyploidy has not been observed in wild uncultivated *Clematis*. *Clematis* is used widely in horticulture and ranks among the most popular ornamental plants, although some *Clematis*, such as North American members of Erickson's (1943) subsection *Integrifoliae* (*C. ochroleuca* Aiton, *C. albicoma* Wherry, *C. viticaulis* Steele, *C. coactilis* (Fernald) Keener, and *C. fremontii* S. Watson) of subg. *Viorna* (Rchb.) Tamura, are rarely used in horticulture (Grey-Wilson 2000).

The edaphic forb, *C. fremontii* is a long-lived, diploid (2n=16) perennial with a distribution that is limited to dolomitic or limestone substrates of calcareous glades and prairies in Missouri and Kansas (Erickson 1943, Gregory 1941). *Clematis fremontii* is of primary

interest because two disjunct populations were recently discovered in Hamilton County, Tennessee and Floyd County, Georgia. If these populations are native rather than introduced, then they effectively represent critically endangered populations in both Tennessee and Georgia, as these are the only two documented occurrences outside of the established range of *C. fremontii*.

In keeping with the array of population-level studies in plants that utilize a phylogeographic approach of sequencing cpDNA and/or nrDNA molecular markers, I sequenced five noncoding cpDNA and the nrDNA (ITS) regions in an attempt to assess the origin of the two disjunct southeastern *C. fremontii* populations as well as to investigate the phylogeography of *C. fremontii* in relation to other taxa in Erickson's (1943) subsection *Integrifoliae* species within the eastern United States: *C. coactilis, C. albicoma, C. viticaulis,* and *C. ochroleuca*. Among a survey of several of the most variable cpDNA regions (sensu Shaw et al. 2005, 2007) few mutations were observed between geographically disparate populations of *C. fremontii* indicating regions within the chloroplast genome are too uninformative to provide phylogeographic information about this species. In a study of 62 accessions of *C. fremontii* using nrDNA ITS sequences Montgomery and Shaw (unpublished data, Part 3 of this thesis) uncovered a staggering number of incompletely sorted cistronic alleles. With so many ITS alleles uncovered, oftentimes more than two alleles per individual and several that are apparently pseudogenes, this gene region is too complex to provide accurate phylogeographic information within this species. Therefore, the origin of the southeastern *C. fremontii* populations remains obscure due to a lack of phylogeographic information in both cpDNA and nrDNA regions, indicating that a region with a quicker divergence time and simpler mode of evolution is necessary to address the original question regarding the origin of the two southeastern *C. fremontii* populations.

The introns of low-copy number nuclear genes have been successful in numerous animal phylogeography studies (Stepian et al. 2001, Crespi et al. 2003, Alstrom et al. 2007), and have now proven useful in low level plant studies as well. Despite the relative novelty of utilizing low-copy nuclear genes in plant phylogeography studies, a variety of low-copy nDNA regions have proved useful across a wide range of taxa (Mort and Crawford 2004). Like cpDNA markers, low-copy number nuclear genes evolve at different rates; therefore careful consideration must be given to the gene chosen for molecular phylogeography research. One

marker in particular, demonstrating a high number of polymorphisms in several intraspecific plant phylogeography studies is Glyceraldehyde 3-phosphate dehydrogenase (*G3pdh*) (e.g. Morrell et al. 2003, Dane et al. 2007, Breen et al. 2009, Tani et al. 2003, Olsen and Schaal 1999). Glyceraldehyde 3-phosphate dehydrogenase encodes the cytosololic enzyme glyceraldehyde 3-phosphate dehydrogenase (GAPDH). Angiosperms contain three forms of GAPDH, a cytosolic form, which is a homotetramer of subunits (GapC) and the marker enzyme of chloroplasts GapAB, which exists in two forms, an A_4 homotetramer or an A_2B_2 heterotetramer (Meyer-Gauen et al. 1994). The GAPDH subunits, GapA, GapB, and GapC, are nuclear encoded (Cerff and Kloppstech 1982). Both the cytosolic and plastid forms of GAPDH catalyse the reversible reduction of 1, 3-diphosphoglycerate to glceraldehyde-3-phosphate; however the cytosolic form catalyses the oxidative step of glycolysis, whereas the chloroplast form catalyses the reductive step of the Calvin cycle (Meyer-Gauen et al. 1994). Sequences of the *G3pdh* gene have been used to study the origin of cultivated Cassava (*Manihot*) (Euphorbiaceae; Olsen and Schaal 1999, Olsen 2002), nucleotide diversity among natural populations of *Populus balmifera* (Salicaceae, Breen et al. 2009), the phylogeography of *Citrullus colocythis* (Cucurbitaceae, Dane et al. 2006), and intraspecific variation in *Cryptomeria japonica* (Cupressaceae; Tani et al. 2003) and in *Hordeum vulgare* ssp. *spontaneum* (Poaceae; Morrell et al. 2003).

The focus of this study was to assess *G3pdh* in *C. fremontii* to determine if this gene region can help illuminate the relationships among midwestern and southeastern populations of this species.

MATERIALS AND METHODS

Plant materials—Twelve populations of *C. fremontii* were sampled; five populations from across the range of the species in Missouri, five populations from across the range of the species in Kansas, and two southeastern populations (one from Tennessee and one from Georgia); these are the only known occurrences in the southeast (Figure 1). Tissue samples consisting of a single leaf were taken from 10 arbitrarily chosen plants at widely spaced intervals throughout each population and stored on ice until DNA extraction. A single whole plant was taken from each population as a voucher specimen to be deposited in the herbarium at the University of Tennessee at Chattanooga (UCHT). Five individuals from one population per species were

sampled in the same manner for *C. albicoma*, *C. viticaulis*, *C. coactilis*, and *C. ochroleuca*. Table 1 lists the sampled populations and their counties and states of occurrence.

DNA extraction, amplification, and sequencing—DNA was extracted from leaves using the DNeasy Plant Mini Kit (Qiagen, Valencia, California, USA). The polymerase chain reaction (PCR) was performed using Eppendorf Mastercycler gradient cycler in 40 µl volumes with the following reaction components: 1 µL template DNA (~ 10-100 ng), 1X ExTaq buffer (PanVera / TaKaRa, Madison, Wisconsin, USA), 200 mmol/L each dNTP, 3.0 mmol/L MgCl2, 0.1 µmol/L each primer, and 1.25 units ExTaq (PanVera / TaKaRa). Reactions included bovine serum albumin at a final concentration of 0.2 mg/mL, which is known to improve amplification from difficult templates. PCR and sequencing primers are listed in Table 2. The PCR protocol described below was preceded by template DNA denaturation at 95°C for 2 minutes and followed by a final extension step of 5 min at 72°C. The PCR cycling conditions were 30 cycles of: denaturation at 95°C for 1 min, primer annealing at 60°C for 1 min, and primer extension at 72°C for 2 min. PCR products were checked on 1% agarose gels before being cleaned with ExoSAP-IT (USB, Cleveland, Ohio, USA). DNA sequencing was performed using both primers. Two internal sequencing primers (Table 2) were used on accessions where sequencing with external primers failed. All DNA sequencing was performed with the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit, v. 3.1 (Perkin-Elmer / Applied Biosystems, Foster City, California, USA) and electrophoresed and detected on an ABI Prism 3100 automated sequencer (University of Tennessee Molecular Biology Resource Facility). Sequencher 4.7 (Gene Codes) was used to edit the DNA strands.

Preliminary alignments and phylogenetic analyses—The partial data set was aligned by eye in MacClade v. 4.0 (Sinauer, Sunderland, Massachusetts, USA). Variable positions in the data matrix were double checked against the original chromatogram files to verify that all base calls were true at all variable positions. In all cases, alignment of potentially informative positions was unambiguous. Indels were coded as binary characters.

A Bayesian analysis of the data was performed using MrBayes 3.1 (Huelsenbeck and Ronquist, 2001) to generate posterior probability distribution using Markov chain Monte Carlo (MCMC) methods. No a priori assumptions about tree topology were made. The statistical

model of DNA substitution, GTR+I+G, was that estimated as the best-fitting maximum likelihood model using MrModeltest 1.1b (Nylander, 2002), which is a simplified version of MODELTEST 3.06 (Posada and Crandall, 1998). The Monte Carlo Markov Chain (MCMC) process was set to run one million generations with four chains. Burn-in was estimated visually by plotting log-likelihood values in Microsoft Excel to determine the number of generations that had run before likelihood values reached an asymptote. To calculate the posterior probability of each bipartition a 50% majority-rule consensus tree was constructed from the remaining trees using PAUP*. *Clematis viorna* was used to root the tree.

To detect interior ancestral haplotypes, TCS v. 1.21 (Clement et al., 2000) was used, which implements a statistical parsimony approach to estimating gene genealogies. TCS was used because it can infer ancestral or intermediate haplotypes (as opposed to assuming that these haplotypes are extinct). For the TCS analysis I used the same sequence alignment that was used in the MrBayes analysis described above, but only *C. fremontii* sequences were included in the analysis.

RESULTS

PCR amplicons of G3pdh—PCR amplicons of the *G3pdh* region were successfully generated for 12 populations of *C. fremontii*, representing ten individuals per population for five populations in Kansas, ten individuals per population for four populations in Missouri and five individuals per one population in Missouri, 22 individuals per one population in Tennessee and 13 individuals per one population in Georgia, and for five individuals from one population per species of *C. coactilis*, *C. viticaulis*, and *C. albicoma* in Virginia; and for one individual of *C. viorna* from North Carolina and one individual of *C. addisonii* from Virginia (Table 1).

Sequencing of G3pdh PCR amplicons—All four primers (Table 2) were used to directly sequence the PCR amplicons of the *G3pdh* region, but many sequences failed due to presumed promiscuous primer binding. Forty-nine individuals were successfully sequenced (highlighted in gray in Table 1), and a list of those individuals and the corresponding primers that generated successful sequence for each individual are also listed in Table 1. Several accessions were partially sequence as a result some primers working (marked with a "1" in Table 1) and other primers failing to work in the sequencing reactions (marked with a "0" in Table 1). Due to

sequencing difficulties, the *G3pdh* dataset is partial (Table 1) as compared to the sampling effort and Part 3, above, and the following results refer to the partial dataset.

***G3pdh in Clematis*—**The sequenced *G3pdh* region in *C. fremontii* ranged from 1565 to 1580 in length, including 268 bp of coding sequence from six exons (four entire and two partial where primers were embedded) (Figure 2). Eighty-three bp located in intron d were excluded from the aligned and coded sequence data due to "messy" peaks on the chromatogram files (Figure 2). Twenty two unique haplotypes were observed in the individuals of *C. fremontii* and one, two, and three haplotypes were uncovered in *C. coactilis*, *C. viticaulis,* and *C. albicoma,* respectively. Ten of the 35 *C. fremontii* individuals sequenced were heterozygous at the *G3pdh* locus. Table 3 reports the observed haplotype and the heterozygosity of each *C. fremontii* accession and the heterozygosity of subsection *Integrifoliae* sequences. The percent heterozygosity in Kansas was 53.8 as compared to 18.2 in the southeast and 18.2 in Missouri (Table 4).

Genealogy of G3pdh haplotypes—Bayesian analyses of 49 *G3pdh* sequences (1362 nt) yielded the phylogenetic tree in Fig. 3. Although resolution of internal branches was mostly low, posterior probabilities were 100 for four clades of three *C. fremontii* individuals from Tennessee, four *C. fremontii* individuals from Georgia, two *C. fremontii* individuals from Kansas, and a clade of seven individuals representing all three subsect. *Integrifoliae* taxa sampled from Virginia (Fig. 2). Each of the four clades with posterior probabilities of 100 are restricted to specific geographic regions. The three clades comprised of *C. fremontii* taxa, represent either Georgia, Tennessee, or Kansas, and the clade comprised of three subsect. *Integrifoliae* taxa is specific to the Alleghany Mountains of Virginia. The large clade with low branch support (56% posterior probability) represents *C. fremontii* individuals from all four regions sampled (Georgia, Tennessee, Missouri, and Kansas).

G3pdh haplotype network—A haplotype network for the *C. fremontii* sequences was reconstructed with the TCS program (Figure 4). Portions of the network were reticulate, as indicated on the haplotype network by closed connectivity of haplotypes. Reticulations in the network indicate possible recombination events among individuals of the same population or homoplasious character states. Haplotypes were grouped into four larger haplotypes (A, B, C, D,

E), and subsequently numbered. Haplotype group A consists of eight haplotypes, all of which occur in a frequency of one, except haplotype A1 which occurs in a frequency of seven. Haplotype group B consists of a single haplotype "B" which occurs in a frequency of one. Haplotype group C consists of three haplotypes that all occur in a frequency of one. Haplotype group D consists of three haplotypes that all occur in a frequency of one. Haplotype group E consists of seven haplotypes. Haplotypes E1, E2, E3, and E4 occur in a frequency of 1. Haplotypes E5, E6, and E7 occur in frequencies of five, four, and two respectively. Haplotypes listed in Table 4 correspond with those identified in Figure 4. Haplotype B, and individual from Kansas is the inferred ancestral haplotype of haplotypes C, D, and E. Haplotype A1 is the inferred ancestral haplotype of the whole network.

DISCUSSION

***G3pdh Sequences*—**Despite the use of two universal *G3pdh* external primers and both forward and reverse internal primers (Table 1), I encountered problematic sequencing, which resulted in failed sequences for 96 out of 144 individuals. Reasons for the sequencing difficulties have yet to be determined, however, low DNA concentration in PCR amplicons, heterozygosity, and promiscuous primer binding are suspected.

G3pdh in Clematis—Although the dataset is partial, a low percentage of heterozygosity (28.6) and numerous nucleotide polymorphisms were encountered in the *C. fremontii* sequences. Populations of *C. fremontii* are isolated with little opportunity for genetic exchange, which befits the low percentage of heterozygosity observed.

***Preliminary phylogenetic analysis*—**The monophyletic grouping of seven individuals representing all three subsect. *Integrifoliae* taxa sampled from Virginia suggests a lack of sequence variation among those species. This lack of variation is surprising because of the high morphological variation between species. Low molecular variance combined with high morphological variation could be indicative of recent radiation. The partitioning of *C. fremontii* sequences into several clades, and the omission of the species from the Virginia clade, suggests this species is more distantly related than the other three subsect. *Integrifoliae* taxa used in this analysis and is possibly an older lineage. The two monophyletic clades, 1) three individuals

from Tennessee and 2) four individuals from Georgia, are strong support for the hypothesis that these populations are relicts and thus display distinct genetic profiles.

Preliminary haplotype network—The haplotype network does not reflect any geographic pattern, with the exception of haplotype C which includes only accessions (3) from Tennessee. The haplotype network represents diploid sequences. Any heterozygous characters were treated as ambiguities in the sequence characterization, and therefore were not counted as character state changes in the TCS analysis. The separation of alleles through cloning would eliminate ambiguities in the dataset due to heterozygosity and would effectively increase the number of potentially informative character state changes. Analyzing both alleles for each individual would increase resolution and could potentially decrease the reticulation found within the network.

LIST OF REFERENCES

Alstrom, P., U. Olsson, P. C. Rasmussen, C. T. Yao, P. G. P. Ericson, and P. Sundberg. 2007. Morphological, vocal and genetic divergence in the *Cettia acanthizoides* comples (Aves: Cettidae). Zoological Journal of the Linnean Society 149(3): 437-452.

Alvarez, I. and J.F. Wendel. 2003. Ribosomal ITS sequences and plant phylogenetic inference. Molecular Phylogenetics and Evolution 29: 417-434.

APG II. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. Botanical Journal of the Linnean Society. 141: 339-436.

Avise, J.C. 2000. Phylogeography: The History and Formation of Species. Harvard University Press, Cambridge, MA.

Baldwin B.G., M.J. Sanderson, J.M. Porter, M.F. Wojciechowski, C.S. Campbell, and M.J. Donoghue. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Ann. Missouri Bot. Gard. 82(2): 247-277.

Breen, A. L., E. Glenn, A. Yeager, M. S. Olson. 2009. Nucleotide diversity among natural populations of a North American poplar (*Populus balsamifera*, Salicaceae) New Phytologist 182 (3): 763-773.

Cerff, R. and K. Kloppstech. 1982. Structural diversity and differential light control of mRNAs coding for angiosperm glyceraldehydes-3-phosphate dehydrogenases. Proc. Natl. Acad. Sci. 79: 7624-7628.

Clement, M., D. Posada, and K. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9: 1657-1659.

Crespi, E. J., L. J. Rissler, and R. A. Browne. 2003. Testing Pleistocene refugia theory: phylogeographical analysis of *Desmognathus wrighti*, a high-elevation salamander in the southern Appalachians. Molecular Ecology 12: 969-984.

Dane, F., J. R. Liu, and C. K. Zhang. 2007. Phylogeography of the bitter apple, *Citrullus colocynthis*. Genetic Resources and Crop Evolution 54(2):327-336.

Erickson, R.O. 1943. Taxonomy of *Clematis* section *Viorna*. Annals of the Missouri Botanical Garden 30(1): 1-62.

Gregory, W.C. 1941. Phylogenetic and Cytological Studies in the Ranunculaceae. Transactions of the American Philosophical Society 31(5): 443-521.

Grey-Wilson. 2000. *Clematis*, the genus: a comprehensive guide for gardeners, horticulturists, and botanists.

Hare, M. P. 2001. Prospects for nuclear gene phylogeography. Trends in Ecology and Evolution 16(12): 700-706.

Meyer-Gauen, G., C. Schnarrenberger, R. Cerff, and W. Martin. 1994. Molecular characterization of a novel, nuclear-encoded, NAD^+ -dependent gleyceraldehyde-3-phosphate dehydrogenase in plastids of the gymnosperm *Pinus sylvestris* L. Plant Molecular Biology 26: 1155-1166.

Morrell, P. L., K. E. Lundy, and M. T. Clegg. 2003. Distinct geographic patterns of genetic diversity are maintained in wild barley (*Hordeum vulagare* spp. *spontaneum*) despite migration. Proceedings of the National Academy of Sciences 100(19): 10812-10817.

Mort, M. E. and D. J. Crawford. 2004. The continuing search: low-copy nuclear sequences for lower-level plant molecular phylogenetic studies. Taxon 53(2): 257-261.

Olsen, K. M. 2002. Population history of *Manihot esculenta* (Euphorbiaceae) inferred from nuclear DNA sequences. Molecular Ecology 11: 901-911.

Olsen, K. M. and B. A. Schaal. 1999. Evidence on the origin of cassava: phylogeography of *Manihot esculenta*. Proceedings of the National Academy of Sciences 96: 5586-5591.

Posoda, D. and K.A. Crandall. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14(9); 817-818.

Sang, T. 2002. Utility of low-copy nuclear gene sequences in plant phylogenetics. Cricital Reviews in Biochemistry and Molecular Biology. 37 (3): 121-147.

Shaw, J., E. B. Lickey, E. E. Schilling, and R. Small. 2007. The tortoise and the hare III: comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms. American Journal of Botany 94:275-288.

Shaw, J. and R.L. Small. 2005. Chloroplast DNA phylogeny and phylogeography of the North American plums (*Prunus* subgenus *Prunus* section *Prunocerasus*, Rosaceae). American Journal of Botany 92(12): 2011-2030.

Small, R. L., J. A. Ryburn, R. C. Cronn, T. Seelanan, and J. F. Wendel. 1998. The tortoise and the hare: choosing between noncoding plastome and nuclear Adh sequences for phylogeny reconstruction in a recently diverged plant group American Journal of Botany 85: 1301-1315.

Small, R. L., R. C. Cronn, and J. F. Wendel. 2004. L. A. S. JOHNSON REVIEW No. 2. Use of nuclear genes for phylogeny reconstruction in plants. Australian Systematic Botany 17: 145–170.

Stepian, C. A., R. H. Rosenblatt, and B. A. Bargmeyer. 2001. Phylogeography of the spotted sand bass, *Paralabrax maculatofaciatus*: Divergence of gulf of California pacific coast populations. Evolution 55(9): 1852-1862.

Tani, N., Y. Tsumura, and H. Sato. 2003. Nuclear gene sequences and DNA variation of *Cryptomeria japonica* samples from the postglacial period. Molecular Ecology 12: 859-868.

Yang T.Y.A.and D. M. Moore. 1999. A revision of the Viorna group of species (section Viorna sensu Prantl) in the genus Clematis (Ranunculaceae). Systematics and Geography of Plants 68:281–303.

Table 1. Taxa used in this investigation, source and voucher numbers, population names, total DNA numbers, corresponding size of PCR amplicons of the G3pdhregion, and corresponding primers used to sequence amplicons. Cells shaded gray contain DNA numbers of complete G3pdh sequences. Successful sequencing reaction with each primer per indiv idual is indicated with a 1, and failed sequenceing reaction is indicated with a O. First letter of population name represents the species; second and third letters represent the state in which the population is located; number distinguishes multiple popluations per state according to order of population sampling.

						Approximate size of *G3pdh* PCR amplicon (kb)	Primer Name			
							GPDX7F	GPDX9R	G3pdh2F	G3pdh2R
Species	Population	Collector; Number; Year	Herb.	State and County	Total DNA #					
C. fremontii S. Watson	FTN1	C. Montgomery; CMo1; 2007	UCHT	USA; TN; Hamilton						
					566	1.5	1	0	1	1
					567	1.5	1	1	1	1
					568	1.5	0	1	0	1
					569	1.5	0	1	1	0
					570	1.5	0	1	1	0
					571	1.5	0	1	1	0
					572	1.5	0	1	1	0
					573	1.5	0	0	0	0
					574	1.5	0	0	0	0
					575	1.5	0	1	0	0
					576	1.5	1	1	1	1
					577	1.5	0	1	1	0
					578	1.5	0	0	0	1
					579	1.5	0	1	1	0
					580	1.5	0	1	0	0
					581	1.5	0	1	0	0
					582	1.5	0	1	0	0
					583	1.5	0	1	0	0
					584	1.5	0	0	0	0
C. fremontii S. Watson	FGA1			USA; GA; Floyd						
					585	1.5	0	0	0	0
					586	1.5	1	0	1	1
					587	1.5	0	0	0	0
					588	1.5	0	1	0	0
					589	1.5	0	0	0	0
					590	1.5	0	0	0	0
					591	1.5	0	1	1	0
					592	1.5	0	1	0	1
					593	1.5	0	0	0	0
					594	1.5	0	0	0	0
					595	1.5	0	1	0	1
					596	1.5	1	1	1	1
					597	1.5	0	1	0	0
C. fremontii S. Watson	FMO1	C. Montgomery; CMo2; 2008	UCHT	USA; MO; Jefferson						
					601	1.5	0	1	1	1
					602	1.5	0	1	1	0
					603	1.5	0	0	0	1
					604	1.5	0	0	0	0
					605	1.5	0	0	1	0
C. fremontii S. Watson	FMO2	C. Montgomery; CMo3; 2008	UCHT	USA; MO; Franklin						
					606	1.5	0	1	1	1
					607	1.5	1	1	1	1
					608	1.5	0	0	1	0
					609	1.5	1	0	1	0
					610	1.5	0	1	0	0
					653	1.5	0	0	1	0
					654	1.5	0	0	1	0
					655	1.5	0	0	1	0
					656	1.5	0	0	1	0
					657	1.5	1	0	1	0

Table 1. Taxa used in this investigation, source and voucher numbers, population names, total DNA numbers, corresponding size of PCR amplicons of the G3pdhregion, and corresponding primers used to sequence amplicons. Cells shaded gray contain DNA numbers of complete G3pdh sequences. Successful sequencing reaction with each primer per indiv idual is indicated with a 1, and failed sequenceing reaction is indicated with a O. First letter of population name represents the species; second and third letters represent the state in which the population is located; number distinguishes multiple popluations per state according to order of population sampling.

						Approximate size of *G3pdh* PCR amplicon (kb)	Primer Name			
							GPDX7F	GPDX9R	G3pdh2F	G3pdh2R
C. fremontii S. Watson	FMO3	C. Montgomery; CMo4; 2008	UCHT	USA; MO; Jefferson						
					611	1.5	0	1	0	0
					612	1.5	1	1	0	0
					613	1.5	1	1	0	1
					614	1.5	0	1	0	1
					615	1.5	0	0	0	0
					658	1.5	0	0	1	0
					659	1.5	0	1	0	1
					660	1.5	0	0	1	0
					661	1.5	0	0	1	0
					662	1.5	0	1	1	0
C. fremontii S. Watson	FMO4	C. Montgomery; CMo5; 2008	UCHT	USA; MO; Washington						
					616	1.5	0	1	0	1
					617	1.5	0	0	0	0
					618	1.5	0	1	1	1
					619	1.5	1	1	1	1
					620	1.5	0	1	0	1
					663	1.5	1	1	1	1
					664	1.5	0	1	1	1
					665	1.5	0	1	1	1
					666	1.5	0	1	1	1
					667	1.5	0	0	0	0
C. fremontii S. Watson	FMO5	C. Montgomery; CMo6; 2008	UCHT	USA; MO; Ozark						
					621	1.5	0	0	1	0
					622	1.5	0	0	0	0
					623	1.5	0	0	0	0
					624	1.5	0	0	0	0
					625	1.5	0	1	0	1
					668	1.5	0	0	0	0
					669	1.5	1	1	1	1
					670	1.5	0	0	0	0
					671	1.5	0	0	0	0
					672	1.5	0	0	1	0
C. fremontii S. Watson	FKS1	C. Montgomery; CMo7; 2008	UCHT	USA; KS; Cloud						
					626	1.5	0	0	0	0
					627	1.5	0	0	0	0
					628	1.5	1	1	0	1
					629	1.5	0	0	1	0
					630	1.5	0	0	1	0
					673	1.5	1	1	1	0
					674	1.5	0	0	1	1
					675	1.5	0	0	0	0
					676	1.5	0	0	0	0
					677	1.5	0	1	0	0
C. fremontii S. Watson	FKS2	C. Montgomery; CMo8; 2008	UCHT	USA; KS; Mitchell						
					631	1.5	0	0	0	0
					632	1.5	0	0	0	0
					633	1.5	0	1	1	1
					634	1.5	0	0	0	0
					635	1.5	1	0	1	0
					678	1.5	0	0	1	0
					679	1.5	1	0	1	1
					680	1.5	1	1	0	1
					681	1.5	0	0	0	0
					682	1.5	0	0	0	0

Table 1. Taxa used in this investigation, source and voucher numbers, population names, total DNA numbers, corresponding size of PCR amplicons of the G3pdhregion, and corresponding primers used to sequence amplicons. Cells shaded gray contain DNA numbers of complete G3pdh sequences. Successful sequencing reaction with each primer per indiv idual is indicated with a 1, and failed sequenceing reaction is indicated with a O. First letter of population name represents the species; second and third letters represent the state in which the population is located; number distinguishes multiple popluations per state according to order of population sampling.

						Approximate size of *G3pdh* PCR amplicon (kb)	Primer Name			
							GPDX7F	GPDX9R	G3pdh2F	G3pdh2R
C. fremontii S. Watson	FKS3	C. Montgomery; CMo9; 2008	UCHT	USA; KS; Rooks						
					636	1.5	1	1	1	0
					637	1.5	1	1	0	1
					638	1.5	0	0	0	0
					639	1.5	0	0	0	1
					640	1.5	0	0	0	0
					683	1.5	1	1	0	1
					684	1.5	0	0	1	0
					685	1.5	0	0	1	0
					686	1.5	0	0	1	0
					687	1.5	0	0	0	0
C. fremontii S. Watson	FKS4	C. Montgomery; CMo10; 2008	UCHT	USA; KS; Phillips						
					641	1.5	0	0	0	0
					642	1.5	0	1	1	0
					643	1.5	0	0	0	0
					644	1.5	0	0	0	0
					645	1.5	0	1	1	0
					688	1.5	0	1	1	0
					689	1.5	0	0	1	0
					690	1.5	0	0	0	0
					691	1.5	0	0	0	0
					692	1.5	1	1	1	1
C. fremontii S. Watson	FKS5	C. Montgomery; CMo11; 2008	UCHT	USA; KS; Jewell						
					646	1.5	0	1	0	0
					647	1.5	0	1	0	0
					648	1.5	1	1	0	0
					649	1.5	1	1	1	1
					650	1.5	0	1	0	0
					693	1.5	0	0	0	0
					694	1.5	0	0	0	0
					695	1.5	1	1	1	1
					696	1.5	1	1	1	1
					697	1.5	1	1	0	0
C. viorna L.	RNC1	C. Montgomery; CMo12; 2008	UCHT	USA; NC; Jackson						
					600	1.5	1	1	1	1
C. addisonii Britton	DVA1	C. Montgomery; CMo16; 2009	UCHT	USA; VA; Montgomery						
					700	1.5	1	1	1	1
C. coactilis (Fernald) Keener	CVA1	C. Montgomery; CMo13; 2009	UCHT	USA; VA; Craig						
					701	1.5	1	1	1	1
					702	1.5	0	1	0	1
					703	1.5	1	1	1	1
					704	1.5	1	1	1	1
					705	1.5	0	0	1	0
C. viticaulis Steele	VVA1	C. Montgomery; CMo14; 2009	UCHT	USA; VA; Bath						
					706	1.5	1	1	1	1
					707	1.5	1	1	1	1
					708	1.5	0	1	0	1
					709	1.5	1	1	1	1
					710	1.5	0	1	1	0
C. albicoma Wherry	AVA1	C. Montgomery; CMo15; 2009	UCHT	USA; VA; Alleghany						
					711	1.5	0	1	1	1
					712	1.5	0	1	1	1
					713	1.5	1	0	1	0
					714	1.5	1	1	1	1
					715	1.5	0	1	1	1

Table 2. Primer names and sequences used to amplify and sequence the *G3pdh* region.

Primer Use	Primer name and sequence (5'-3')
PCR amplification and sequencing	**GPDX7F:** GAT AGA TTT GGA ATT GTT GAG G (Strand et al., 1997) **GPDX9R:** AAG CAA TTC CAG CCT TGG (Strand et al., 1997)
Internal sequencing	**G3pdh2F:** CTA ATG GAA ACT TAC TGG **G3pdh2R:** AAC TAA CGA TAG AAC TAG AC

Table 3. Taxa used in this investigation, population names, DNA numbers, observed G3pdh haplotype, and heterozygosity or homozygosity of G3pdh region. G3pdh haplotypes are assigned to sequences of C. fremontii accessions only. Sequences displaying heterozygosity or homozygosity have been marked with an asterisk in the appropriate column.

Species	Population	Successful *G3pdh* Sequence (DNA#)	*G3pdh* Haplotypes	Heterozygous	Homozygous
C. fremontii S. Watson	FTN1				
		566	C1		*
		567	C3		*
		570	A5		*
		571	A6	*	
		576	C2		*
		579	D2		*
C. fremontii S. Watson	FGA1				
		586	E6		*
		591	E6		*
		592	E6		*
		595	A1	*	
		596	E6		*
C. fremontii S. Watson	FMO1				
		601	A1		*
		602	A8	*	
C. fremontii S. Watson	FMO2				
		606	A1		*
		607	E4		*
C. fremontii S. Watson	FMO3				
		613	A4		*
C. fremontii S. Watson	FMO4				
		618	A3		*
		619	A1		*
		663	A1		*
		664	A1	*	
		666	A2		*
C. fremontii S. Watson	FMO5				
		669	A1		*
C. fremontii S. Watson	FKS1				
		628	E5		*
		674	E2	*	
C. fremontii S. Watson	FKS2				
		633	E5		*
		679	E5	*	
		680	E5	*	

Table 3. Taxa used in this investigation, population names, DNA numbers, observed G3pdh haplotype, and heterozygosity or homozygosity of G3pdh region. G3pdh haplotypes are assigned to sequences of C. fremontii accessions only. Sequences displaying heterozygosity or homozygosity have been marked with an asterisk in the appropriate column.

Species	Population	Successful *G3pdh* Sequence (DNA#)	*G3pdh* Haplotypes	Heterozygous	Homozygous
C. fremontii S. Watson	FKS3				
		637	B	*	
		683	E5	*	
C. fremontii S. Watson	FKS4				
		645	D3		*
		688	E1	*	
		692	A7		*
C. fremontii S. Watson	FKS5				
		649	E7		*
		695	E7		*
		696	D1	*	
C. viorna L.	RNC1				
		600		*	
C. addisonii Britton	DVA1				
		700			*
C. coactilis (Fernald) Keener	CVA1				
		701		*	
		703		*	
		704		*	
C. viticaulis Steele	VVA1				
		706			*
		707			*
		709			*
		710			*
C. albicoma Wherry	AVA1				
		711		*	
		712		*	
		714		*	
		715		*	

Table 4. Percent heterozygosity and frequency of heterozygous and homozygous *G3pdh* sequences obtained for accessions of *Clematis fremontii* populations across three geographic regions.

Region (populations)	Heterozygous (n)	Homozygous (n)	% Heterozygosity
Southeastern Glade (FTN1, FGA1)	2	9	18.2
Midwestern Glade (FMO1, FMO2, FMO3, FMO4, FMO5)	2	9	18.2
Midwestern Prairie (FKS1, FKS2, FKS3, FKS4, FKS5)	7	6	53.8

Figure 1. Map of *C. fremontii* and subsection *Integrifoliae* populations and subsect. *Viorna* outgroup taxa sampled in this study.

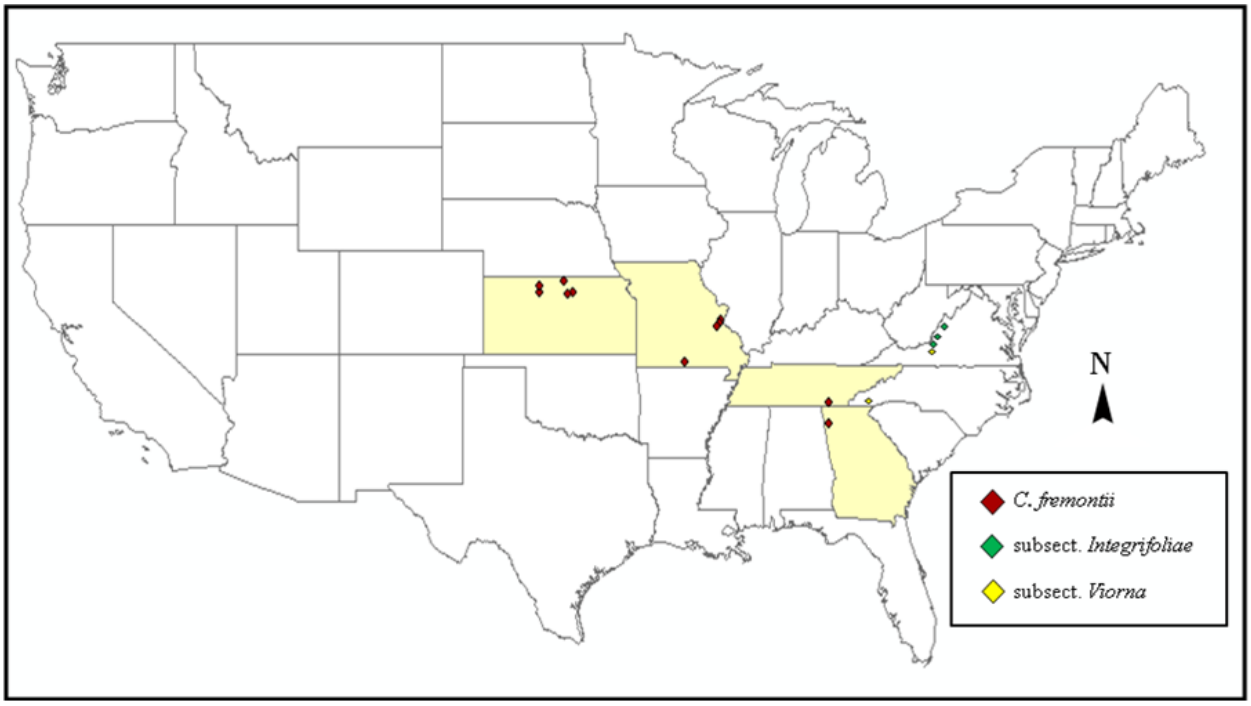


Figure 2. Sequenced region (bp) of glyceraldehydes 3-phosphate dehydrogenase (*G3pdh*). Primer binding sites are indicated by arrows. Exon regions are B, C, D, and E. Introns are a, b, c, and d. Open region in intron d represents 83 bp of ambiguous sequence that was deleted from the aligned and coded dataset.

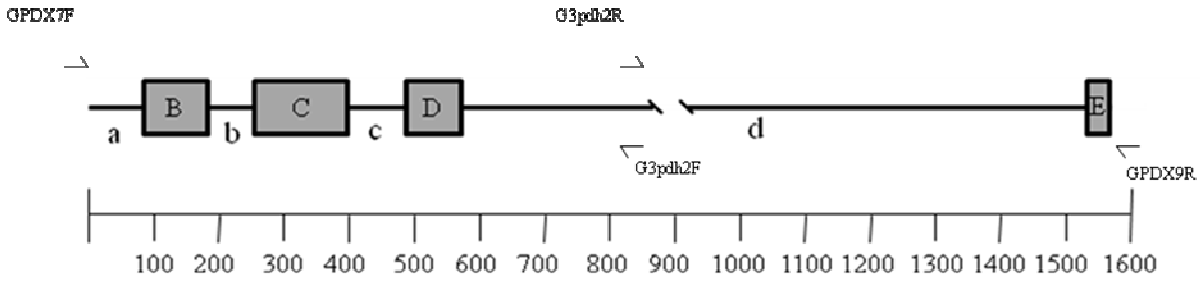


Figure 3. Bayesian analysis of aligned *G3pdh* sequences. Numbers reported represent posterior probabilities of branches. To the right of each clade, species included within that clade are listed and the state in which accessions of *C. fremontii* were collected is listed in parentheses (e.g. *C. fremontii* (TN))

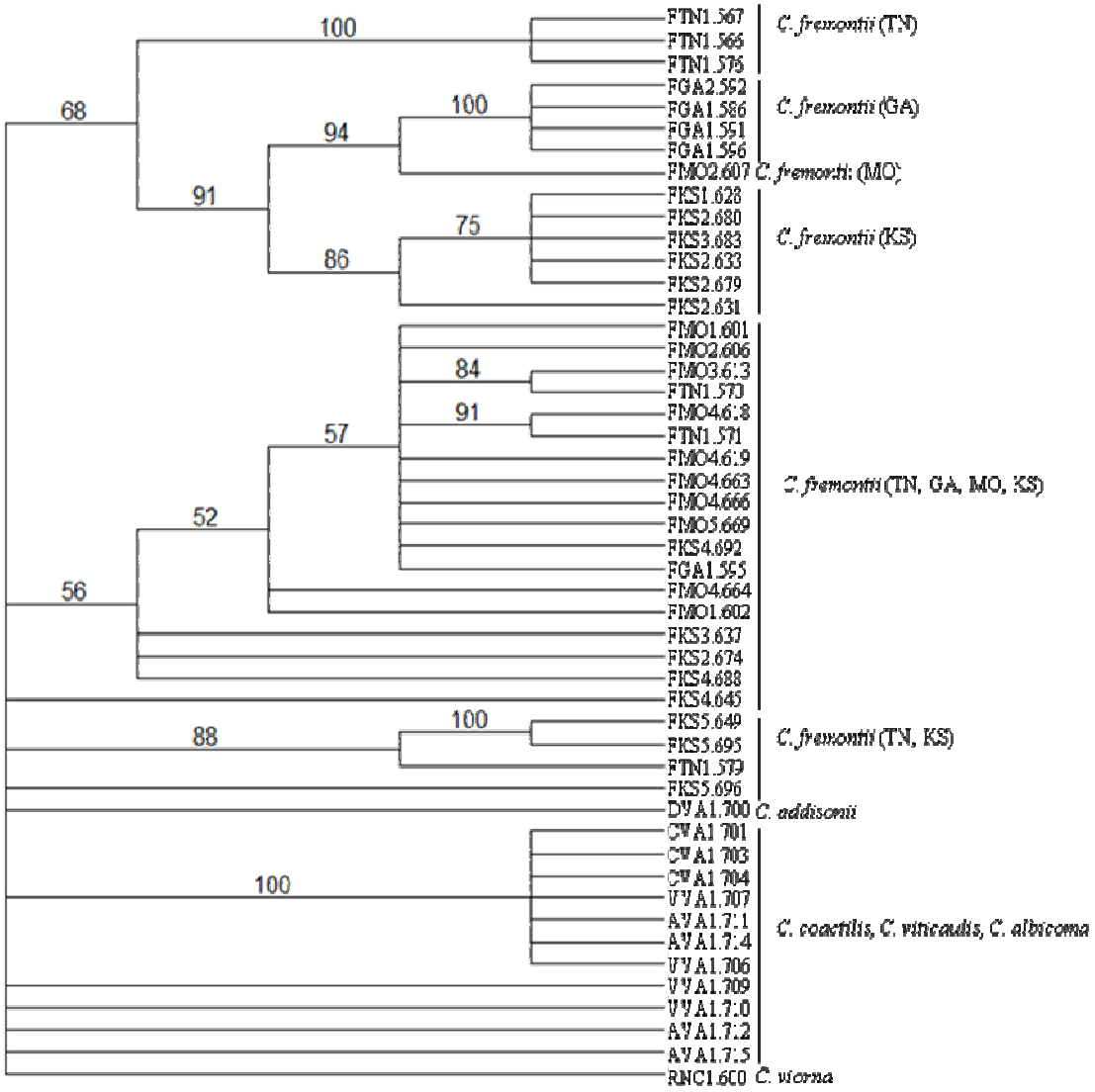
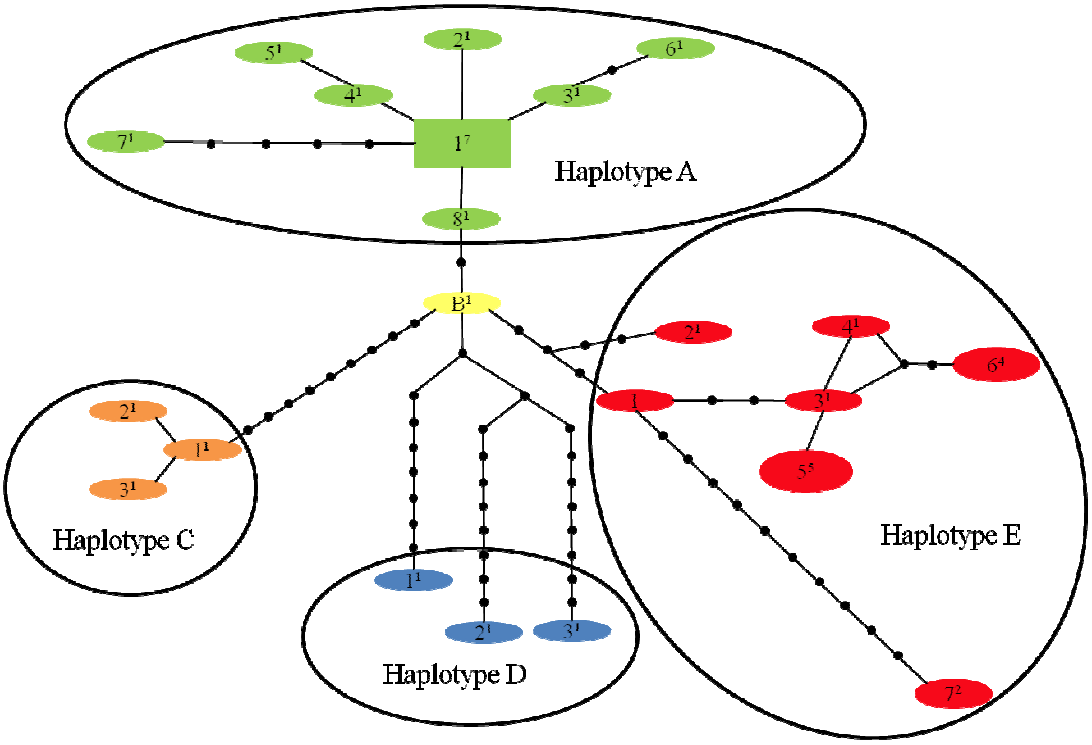


Figure 4. Haplotype network of successful *G3pdh* sequences for individuals of *Clematis fremontii*. Ovals represent sampled haplotypes with relative size indicating frequency of observation and with a rectangle indicating the highest frequency of observation in the haplotype network. Solid circles represent unobserved haplotypes with inferred mutational steps, and solid lines connecting haplotypes represent single nucleotide polymorphisms or indel mutations. Haplotypes are assigned haplotype groups by letter (A-E), boundaries of each haplotype group are indicated with a black outline. Each individual haplotype within a group is assigned a number; so haplotypes are named accordingly (e.g. A1). Superscripts after the haplotype number represent haplotype frequency.



Part 5: **General Conclusions, a Focused Conservation Effort, and Future Directions**

General Conclusions

Clematis fremontii, and the other species in subsection *Integrifoliae* likely represent recently diverged species. This suggestion is strengthened by a broad phylogenetic study of *Clematis* in which Miikeda et al. (2006) concluded that low genetic variation coupled with high morphological diversity within the genus likely indicates a recent rapid radiation. My data support this hypothesis in that I uncovered 333 unique ITS copies in only 86 accessions, representing eight species of *Clematis* subgenus *Viorna*, including 5 species of subsection *Integrifoliae*. Such a high frequency of unique copies indicates concerted evolution of nrDNA repeats has not yet occurred to completion. This incomplete concerted evolution is likely a result of inadequate time for completion due to the putative recent divergence of species within subsection *Integrifoliae* and likely within the entire subgenus *Viorna*. Therefore, while the data are equivocal in terms of answering the question of whether the eastern populations are native, the unique ITS copies discovered in the Tennessee population and unique G3pdh haplotypes identified in the Georgia population as well as in the Tennessee population are interesting and suggestive of natural populations.

A Focused Conservation Effort

As mentioned in Part 1 of this thesis, the Tennessee population of *C. fremontii* is in danger of being extirpated because it resides on commercially zoned private property that is currently for sale. In 2007 Dr. J. Shaw contacted the landowner, Mr. Lewis Card, who explained that he was not interested in selling the land to the state (since typically the state "could not give what the property is worth"). To my knowledge, the land has not been sold because the predominance of shallow limestone bedrock would make building on this property relatively costly due to the blasting required for the installation of underground pipes and wires.

I have been in contact with Todd Crabtree, the Tennessee Natural Heritage State Botanist, and he has strong interest in relocating the Tennessee population to Yuchi Glade in Rhea County, Tennessee. While I could not unequivocally determine if the eastern populations of *C. fremontii* were native populations, I did uncover much genetic variation in the Tennessee population. Since each of the individuals that I sampled from the Tennessee population were marked with GPS coordinates, it is possible to use these geographical points in conjunction with

the allelic information uncovered in this investigation to ensure that the majority of the genetic variation in the Tennessee population is captured in the relocation effort.

FUTURE DIRECTIONS

The future of the eastern populations is dubious. The Georgia population resides on protected land on the Berry College campus in Rome, Georgia; however, the population at this site appears to be in danger of being completely overgrown by pine trees. Additionally, as mentioned above, the Tennessee population is not protected. This project is worth continuing for these reasons alone and the data uncovered here, especially the preliminary *G3pdh* data, show that there is hope for deciphering the phylogeographic relationships of this beautiful, interesting, and thought-provoking species.

Future work on this project will likely involve cloning. Cloning would strengthen the dataset in several ways. The current data cannot distinguish whether *G3pdh* is single copy in *Clematis fremontii*, however a lack of minor peaks "shadows" under the major peaks in the chromatogram files indicates the gene is likely single copy. Comparing PCR amplicons to the original *G3pdh* sequence would help confirm or deny the presumption that *G3pdh* is single copy in *Clematis fremontii*. Additionally, cloning offers the advantage of separating parental alleles, which parses out ambiguous nucleotide positions from double peaks in the direct sequence of heterozygous individuals and also doubles the sample size of a dataset by providing two sequences per individual (one per each allele) rather than one. Separation of alleles would thereby increase the number of informative characters in double peak positions of the original sequences and would provide double the parental lineage information per individual in heterozygotes. Doubling the sample number through cloning and increasing the number of informative characters could help resolve some of the confounding relationships found in the current analyses, such as reticulation within the haplotype network and the dubious geographic relationships among clades.

In addition to cloning, some increased sampling of other section *Integrifoliae* species is needed. *Clematis fremontii* has been heavily sampled, and the current sampling is representative of the entire range of the species. Only one population per species was sampled for *C. coactilis*, *C. viticaulis* and *C. albicoma*. These species are narrowly restricted to only a few populations within a few counties of Virginia and West Virginia (*C. albicoma*), and therefore the current

sampling is an adequate representation of the distribution of these species for the purposes of a phylogeographic study of *C. fremontii* in the Eastern United States. In contrast to the other subsect. *Integrifoliae* taxa which have narrow distributions, *Clematis ochroleuca* has a broad distribution that spans from Georgia to Staten Island, NY. Previous taxonomic treatments link *C. ochroleuca* to other subsect. *Integrifoliae* taxa. Prior to Erickson's treatment (1943), both *C. albicoma* and *C. fremontii* had been included within *C. ochroleuca* (Gray 1895, Small 1933, James 1883). Additionally, *C. ochroleuca* and *C. albicoma* have been suggested as the progenitors of the putative hybrid *C. coactilis* (Keener 1967, 1975). To date only one population has been collected in the southern tip (GA) of the species range. A more robust sampling of *C. ochroleuca* from across the entire species range is necessary to better represent the diversity of *G3pdh* within *C. ochroleuca* and to assess the relatedness of *C. ochroleuca* to *C. fremontii* and other subsect. *Integrifoliae* species.

Another necessity in phylogeographic study is outgroup taxa of the appropriate ancestral distance in order to polarize the dataset. The two outgroup taxa in the preliminary *G3pdh* study are species within subsection *Viorna* (*C. viorna* and *C. addisonii*). Subsections *Integrifoliae* and *Viorna* are included in subgenus *Viorna*. *Clematis viorna* and *C. addisonii* did not segregate from the subsect. *Integrifoliae* taxa in the G3pdh Bayesian tree nor in the ITS Bayesian tree, which indicates species within section *Viorna* are perhaps too closely related to *C. fremontii* for use as outgroup taxa. I recommend including several other subgenus *Viorna* species in addition to a few distantly related outgroup taxa that will serve to better polarize the dataset.

LIST OF REFERENCES

Erickson, R.O. 1943. Taxonomy of *Clematis* section *Viorna*. Annals of the Missouri Botanical Garden 30(1): 1-62.

Gray, A. 1892. "Ranunculaceae," pp. 1-57. In: B.L. Robinson (ed.), Synoptical flora of North America, Vol. I, pt. 1. American Book Co., New York.

James, J.F. 1883. A revision of the genus *Clematis* of the United States; embracing descriptions of all the species, their systematic arrangement, geographical distribution, and synonomy. Journal of Cincinnati Society of Natural History 6: 118-135.

Keener, C. S. 1967. A biosystematic study of Clematis subsection Integrifoliae (Ranunculaceae). J. Elisha Mitchell Sci. Soc. 83: 1-40.

Keener, C. S. 1975. Studies in the Ranunculaceae of the Southeastern United States. I. *Anemone* L. Castanea 40(1) 36-44.

Miikeda, O., K. Kita, T. Handa, T. Yukawa. 2006. Phylogenetic relationships of *Clematis* (Ranunculaceae) based on chloroplast and nuclear DNA sequencesBotanical Journal of the Linnean Society 152: 153–168.

Small, J.K. 1933. Manual of the southeastern flora. Publ. by the author, New York.

Curriculum Vitae

Meredith Montgomery
Department of Biological and Environmental Sciences
615 McCallie Avenue
The University of Tennessee, Chattanooga
Chattanooga, TN 37403-2598

Phone: 423-425-4341
email:meredith-montgomery@utc.edu

EDUCATION

M.S. Environmental Science (Molecular Systematic/ Genetics Concentration). Candidate. University of Tennessee at Chattanooga.

Incomplete concerted evolution in the non-hybrid diploid Clematis fremontii S. Watson (Ranunculaceae).

B.S. Environmental Science. 1998. University of Tennessee at Chattanooga; Magna Cum Laude

ACADEMIC/RESEARCH APPOINTMENTS

2007 – 2009: Teaching Assistant, Department of Biological and Environmental Sciences

GRANTS AND AWARDS

2007 Received the Marie Mellinger Research Grant, a competitive award from the Georgia Botanical Society, to fund my research on *Clematis fremontii* - $1500.

2007 Received Student Research Award, a competitive award from the Geospatial Information and Technology Association to fund my research on *Clematis fremontii* - $500.

2008 Received Provost Student Award to generate DNA sequences from Missouri and Kansas populations of my study species - $1000.

2008 Received Student Research Award, a competitive award from the Geospatial Information and Technology Association to fund my research on *Clematis fremontii* - $1000.

2008 1st place for "Outstanding Presentation" at Tennessee Academy of Sciences.

2009 Graduate student association travel award to attend ASB in Birmingham, AL.

2009 ASB travel award to attend the 2009 meeting in Birmingham, AL.

2009 Sigma Xi Graduate Student Research Award at UTC

PROFESSIONAL PRESENTATIONS

MONTGOMERY, M. AND J. SHAW. 2009. Phylogeography of *Clematis fremontii* S. Wats. (Ranunculaceae) in eastern North America. Paper presented by J. Shaw at the Botanical Society of America meeting in Snowbird, UT.

MONTGOMERY, M. AND J. SHAW. 2009. Phylogeography of *Clematis fremontii* S. Wats. (Ranunculaceae) in eastern North America. Paper presented at the Association of Southeastern Biologists meeting in Birmingham, AL.

MONTGOMERY, M. AND J. SHAW. 2008. *Clematis fremontii* (Ranunculaceae) in Tennessee and Georgia: recent introductions or relict populations? Paper presented at the Tennessee Academy of Sciences meeting in Nashville, TN.

MONTGOMERY, M. AND J. SHAW. 2008. *Clematis fremontii* (Ranunculaceae) in Tennessee and Georgia: recent introductions or relict populations? Poster presented at the Society for Conservation Biology meeting in Chattanooga, TN.

MONTGOMERY, M. AND J. SHAW. 2007. *Clematis fremontii* (Ranunculaceae) in Tennessee and Georgia: recent introductions or relict populations? Poster presented at the Association of Southeastern Biologists meeting in Spartanburg, SC.

PROFESSIONAL MEMBERSHIPS

Association of Southeastern Biologists
Botanical Society of America
Geospatial Information and Technology Association
Tennessee Academy of Sciences

TEACHING/MENTORING EXPERIENCE

University of Tennessee at Chattanooga

Undergraduate Princ. of Biology (BIOL 121) laboratory (3 sections of 24 students per semester, for 5 semesters)

Guest Lecture. Molecular methods and plant conservation genetics. ESC 480 Seminar on Environment

Mentored undergraduate student Clea Klagstad, for two semesters in molecular lab techniques

Undergraduate laboratory section 500 of BIOL 207 Plant Morphology, undergraduate teaching assistant

PREVIOUS WORK EXPERIENCES/SERVICE

2009. Medical Records Reconciliation Analyst. Medibase Corp. Johns Hopkins Hospital, Columbia, MD.

2009. Reading Mentor. United Way summer reading program. Chattanooga, TN

2008. Guest Judge. Science Fair, grades 3-5. McConnell Elementary, Hixson, TN

2004-2009. Assistant Manager. Zumfoot shoes, formerly Birkenstock. Chattanooga, TN

2002-2004. Sales Associate. Birkenstock. Chattanooga, TN

2002-2006. Captain of Signal Mountain East Precinct Democratic Voting Committee